

## Tätigkeitsbericht 2019 der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ)

Veröffentlichungsversion / Published Version  
Tätigkeitsbericht, Jahresbericht / annual report

*Diese Arbeit wurde durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert (Förderkennzeichen: 01UW1802).*

### Empfohlene Zitierung / Suggested Citation:

Rat für Sozial- und Wirtschaftsdaten (RatSWD). (2020). *Tätigkeitsbericht 2019 der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ)*. Berlin. <https://doi.org/10.17620/02671.56>

### Nutzungsbedingungen:

*Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:*  
<https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

*This document is made available under a CC BY Licence (Attribution). For more Information see:*  
<https://creativecommons.org/licenses/by/4.0>

# Tätigkeitsbericht 2019

der vom RatSWD akkreditierten  
Forschungsdatenzentren (FDZ)



GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung



Rat für Sozial- und Wirtschaftsdaten (RatSWD)

# Tätigkeitsbericht 2019

der vom RatSWD akkreditierten  
Forschungsdatenzentren (FDZ)



## Inhaltsverzeichnis

Vorwort der Vorsitzenden des FDI Ausschusses .....	6
<b>1 Überblick über die Forschungsdateninfrastruktur (FDI)</b>	
<b>des Rates für Sozial- und Wirtschaftsdaten (RatSWD)</b> .....	8
Aktuelle Kennzahlen .....	8
Ausgangspunkt für die Etablierung des Netzwerks aus Forschungsdatenzentren .....	8
Weiterentwicklung und Sicherung der Forschungsdateninfrastruktur .....	9
Akkreditierungsprozess und Qualitätssicherung .....	11
Innovationspotential für die Forschung heben .....	11
Neue Akkreditierungen .....	12
<b>2 Struktur der Forschungsdatenzentren (FDZ)</b> .....	15
Personal .....	15
Kooperationen und Forschungsaktivitäten .....	16
Wissenschaftliche Publikationen .....	18
<b>3 Archivierung und Qualitätssicherung</b> .....	19
Konzepte für die langfristige Verfügbarmachung der Daten .....	19
Archivierung .....	20
Qualitätssicherung der Datensätze .....	20
<b>4 Datenangebot und Datennutzung</b> .....	24
Umfang des Datenangebotes .....	24
Zeitpunkt der Verfügbarmachung der Datensätze und Gebühren .....	24
Forschungsbudget auf Basis der bereitgestellten Forschungsdaten .....	26
Nutzungsumfang .....	27
<b>5 Service für Nutzerinnen und Nutzer</b> .....	30
Qualifizierungsmaßnahmen .....	31
Qualitätssicherung der Serviceangebote .....	32
Zugangswege und Formate für die Datenbereitstellung .....	32
Bereitstellung von Arbeitshilfen .....	34
Zeitdauer von der Antragstellung bis zum Datenzugang .....	35
Bewerbung der Daten .....	37
<b>6 Aktuelle Entwicklungen der Forschungsdateninfrastruktur</b>	
<b>in den Sozial-, Verhaltens- und Wirtschaftswissenschaften</b> .....	38
Internationalisierung .....	38
Innovationen und Neuerungen zur bestehenden Dateninfrastruktur .....	40
Weiterentwicklung der Infrastruktur .....	40
<b>7 Fokusthema Datenzusammenführung</b> .....	41
Bedeutung und Potential des Themas Datenverknüpfung für die FDZ .....	41
Ebenen der Datenverknüpfung .....	43
Zugangswege für Verknüpfungen .....	44
<b>8 Beschwerdemanagement</b> .....	46
Aktuelle Beschwerdeverfahren des Berichtsjahres 2018 .....	46
<b>Anhang</b>	
Anhang A: Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD .....	48
Anhang B: Index und Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD .....	52
Anhang C: Die Monitoringkommission .....	58
Anhang D: Mitwirkende bei der Erstellung des Tätigkeitsberichts 2019 .....	59

## Vorwort der Vorsitzenden des FDI Ausschusses

Den Zugang zu Forschungsdaten vereinfachen, Datengewinnung, -dokumentation und -speicherung methodisch weiterentwickeln und die Forschungsdateninfrastruktur koordinieren – mit diesen Zielen vernetzen sich die Forschungsdatenzentren (FDZ) im Ständigen Ausschuss Forschungsdateninfrastruktur (FDI Ausschuss). Wir freuen uns, dass sich das Netzwerk von 34 FDZ im Jahr 2019 inzwischen im Frühjahr/Sommer 2020 um vier neue FDZ erweitert hat und damit zusätzliche Datenangebote für die Migrations-, Gesundheits- und Finanzmarktforschung aber auch speziell zu raumbezogenen Hintergrundfaktoren zur Verfügung stehen. Dieser Erfolg baut auf Vorarbeiten im für diesen Tätigkeitsbericht relevanten Berichtsjahr 2019 auf. Auch für die Zukunft möchten wir weitere Institutionen dazu einladen, durch ihre Akkreditierung das Datenangebot für die Sozial-, Verhaltens- und Wirtschaftswissenschaften zu erweitern und im Netzwerk des Rates für Sozial- und Wirtschaftsdaten (RatSWD) an der Verbesserung der Dateninfrastruktur mitzuwirken.

### KonsortSWD: Beispiele für Angebote und Maßnahmen von und für FDZ im Bereich „Datenzugang“

- Unterstützung bei der Zertifizierung mit einem Core Trust Seal oder der Beratung neuer FDZ bei der RatSWD-Akkreditierung
- Erweiterung eines Netzwerks von miteinander verbundenen sicheren Gastwissenschaftsarbeitsplätzen
- Aufbau einer föderierten Archivierungsinfrastruktur für Daten aus der qualitativen Sozialforschung
- Etablierung eines Trainingsprogramms zur Verbesserung der Datenmanagementfähigkeiten
- Betreuung eines moderierten Online-Foren für den Dialog zwischen Datenproduzierenden und Datennutzenden

Der FDI Ausschuss erarbeitete in den letzten Jahren gemeinsame Entwicklungsziele. Diese flossen unmittelbar in die Planungen des Konsortiums für Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (KonsortSWD) ein. KonsortSWD wird im Rahmen der von der Gemeinsamen Wissenschaftskonferenz (GWK) initiierten Nationalen Forschungsdateninfrastruktur (NFDI) ab Oktober 2020 vorerst für fünf Jahre finanziert – ein großer Erfolg unter anderem auch von RatSWD und FDI Ausschuss. KonsortSWD wird als breites Netzwerk von Universitäten, Forschungseinrichtungen und insbesondere von FDZ als digitaler und vernetzter Wissensspeicher dienen. Davon werden in erster Linie Forschende als auch die FDZ selbst profitieren. Die Förderung im Rahmen der NFDI ermöglicht es, die erfolgreiche Arbeit des FDI Ausschusses in Kooperation mit dem RatSWD auch über das Jahr 2020 weiterzuführen.

Über den FDI Ausschuss hinaus vernetzen sich mittlerweile außerdem 17 FDZ in einem Netzwerk, das sich die Verbesserung der Zugänglichkeit von Forschungsdaten und ein qualitätsgesichertes Forschungsdatenmanagement der Sozial-, Verhaltens- und Wirtschaftswissenschaften zur Aufgabe gemacht hat (Economic and Social Sciences goInG FAIR Implementation Network (EcoSoc-IN)). EcoSoc-IN wurde durch den RatSWD gegründet und ist offen für die Beteiligung anderer nationaler und internationaler Einrichtungen und Personen (auch ohne akkreditierte Forschungsdatenzentren). Das Ziel von EcoSoc-IN ist die Umsetzung der FAIR Prinzipien für die Sozial-, Verhaltens- und Wirtschaftswissenschaften. Die FAIR Prinzipien zielen darauf, Daten digital zugänglich zu machen, indem die Auffindbarkeit (Findability), Zugänglichkeit (Accessibility), Interoperabilität (Interoperability) und Nachnutzbarkeit (Reusability) von Forschungsdaten für Menschen und insbesondere Maschinen verbessert wird. EcoSoc-IN forciert den Austausch über die weitere Verbesserung des Zugangs zu sensiblen Daten in den Sozial-, Verhaltens- und Wirtschaftswissenschaften und verwandten Disziplinen und auf internationaler Ebene. Vor allem geht es darum, die Zugänge und Verfahren verstärkt zu standardisieren. Das Netzwerk unterstützt also nutzungsfreundliche Lösungen für den Zugang zu sensiblen Daten und ihrer Verknüpfung und adressiert damit die speziellen datenschutzrechtlichen Bedarfe der Sozial-, Verhaltens- und Wirtschaftswissenschaften. Zur Unterstützung der FDZ initiierten Mitglieder von EcoSoc-IN im Jahr 2019 ein White Paper zur Umsetzung der FAIR Kriterien, das in Kürze erscheinen wird.

Der vorliegende Tätigkeitsbericht 2019 der vom RatSWD akkreditierten FDZ zeichnet ein Bild von der Forschungsdateninfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften. Die Herausforderung in den RatSWD-nahen Disziplinen besteht häufig darin, Daten, die nicht immer für die Wissenschaft erhoben wurden, für Forschende zugänglich zu machen und gleichzeitig hinreichende datenschutzrechtliche und forschungsethische Rahmenbedingungen zu schaffen. Die Vielzahl der methodischen Ansätze, Datentypen und -inhalte erfordert dabei eine große Spannbreite von Dokumentationsansätzen oder Anonymisierungs- bzw. Pseudonymisierungsschritten. Innerhalb der forschungsbasierten Infrastruktur der FDZ werden hierfür Standards gesetzt, damit der Forschung analysierbare und gut dokumentierte Daten zur Verfügung stehen. Gleichzeitig nimmt die Vielfalt von Datentypen mit jeweils spezifischen Kuratierungsbedarfen beständig zu.

Mit dem jährlichen Tätigkeitsbericht überprüfen wir die Einhaltung der gemeinsam erarbeiteten Standards, sichern so die Qualität des Angebots und geben einen Überblick über die Tätigkeiten der durch den RatSWD akkreditierten FDZ. Dazu liefert der Bericht Antworten auf wichtige Fragen wie: Welche Angebote richten FDZ an Forschende? Wie wird dieses Angebot genutzt? Wie stellen die FDZ Daten und Metadaten, die die Datensätze beschreiben, nachhaltig bereit und wie wird deren Qualität gesichert? Wie viele Personen arbeiten im Schnitt in einem FDZ und welche Arbeitsschritte fallen an? Oder: Welche FDZ mit welchem Datenangebot gibt es derzeit eigentlich?

Das diesjährige Schwerpunktkapitel „Datenzusammenführung“ gibt einen Überblick dazu, inwiefern Daten aus verschiedenen Datenerhebungen miteinander verknüpft werden können oder auch bereits verknüpft wurden, so dass Analysen auf einer noch breiteren Datenbasis stehen und auch neue Fragestellungen beantwortet werden können. Die Ergebnisse zeigen, dass die Daten der FDZ hierfür großes Potential bergen dabei aber forschungsethische und datenschutzrechtliche Erwägungen beachtet werden müssen.

Der RatSWD und der FDI Ausschuss danken den Mitgliedern der Monitoringkommission und dem Team aus der Geschäftsstelle sehr herzlich für ihren persönlichen Einsatz bei der vertrauensvollen Ausarbeitung des Tätigkeitsberichts 2019. Mit der ständigen Weiterentwicklung der Qualitätssicherung trägt die Monitoringkommission wesentlich dazu bei, die Forschungsdateninfrastruktur sowie deren Leistungsfähigkeit detailliert darzustellen und Handlungsempfehlungen auszuarbeiten. Die Nutzungszahlen zeigen, dass die Nachfrage nach qualitativ hochwertigen Daten weiter wächst. Moderne Dateninfrastrukturen werden immer mehr zu einem unverzichtbaren Baustein erfolgreicher Forschung!



Dr. Pascal Siegers  
(Vorsitzender des FDI Ausschusses)



Dr. Jan Goebel  
(stellv. Vorsitzender des FDI Ausschusses)



# 1 Überblick über die Forschungsdateninfrastruktur (FDI) des Rates für Sozial- und Wirtschaftsdaten (RatSWD)



**294**

Vollzeit-  
äquivalente



**521**

Wissenschaftliche  
Publikationen  
des FDZ Personals



**4.371**

Datensätze  
davon 431  
neue Datensätze



**55.270**

Externe  
Datennutzende

Der Ausschuss für Forschungsdateninfrastruktur (FDI Ausschuss) ist ein dynamisches dezentrales Netzwerk mit inzwischen 38 vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ), von denen vier im Jahr 2020 neu akkreditiert wurden (Stand Juli 2020) – die im Folgenden berichteten Zahlen stammen von den 2019 akkreditierten 34 FDZ.<sup>1</sup>

In den FDZ werden Daten archiviert und über verschiedene Zugangswege unter Einhaltung des Datenschutzes für die Wissenschaft zugänglich gemacht. Das Spektrum umfasst dabei etwa Studien- daten aus den verschiedenen Disziplinen der Sozial-, Verhaltens- und Wirtschaftswissenschaften ebenso wie Geo- und Raumdaten, Finanzdaten, statistische Daten von Bund und Ländern sowie Register- und Sozialversicherungsdaten. Nicht nur quantitative Daten, sondern auch qualitative Daten sind inzwischen zugänglich. Das Netzwerk wird beständig um weitere FDZ erweitert und damit auch um weitere Erhebungsmethoden sowie neue Datentypen und -formate angereichert. Der stetige Zuwachs an FDZ und steigende Nutzungszahlen deuten darauf hin, dass das Modell zukunftsfähig ist und eine Akkreditierung durch den RatSWD gegenüber Drittmittelgebenden und Datennutzenden als Qualitätssiegel gilt und die empirische Forschung erleichtert.

## Aktuelle Kennzahlen

Die Kennzahlen für das Jahr 2019 unterstreichen diese erfolgreiche Entwicklung der FDZ-Landschaft: Zum Stichtag 31.12.2019 waren in den FDZ insgesamt 294 Vollzeitäquivalente (VZÄ) beschäftigt, davon etwa zwei Drittel als wissenschaftliches Personal. Das FDZ-Personal verfasste 2019 insgesamt 521 wissenschaftliche Publikationen. Eine positive Zahl in vielerlei Hinsicht: Die (Nach-)Nutzung der eigenen Forschungsdaten sichert den Anschluss an die Forschungscommunity, fördert beim FDZ-Personal die Beratungskompetenz, dient als Qualitätskontrolle und steigert zugleich die Bekanntheit der Datensätze. Mindestens 2.359 wissenschaftliche Publikationen beruhen ganz oder teilweise auf den 4.371 von den FDZ bereitgestellten Datensätzen.

Für die Datennutzenden wurden 2019 insgesamt 431 neue Datensätze bereitgestellt. Der Zuwachs fällt damit um 62 Datensätze größer aus als im Vorjahr – bei gleicher Anzahl der FDZ. Das Nachnutzungspotenzial von Forschungsdaten wächst. Von diesem breiten Angebot machten die Datennutzenden 2019 ausgiebig Gebrauch: Es gab insgesamt 55.270 Datennutzende in den FDZ, d. h. im Vergleich zum Vorjahr sind mehr als 8.600 Datennutzende hinzugekommen. Die FDZ konnten 79.878 freie Downloads verzeichnen, etwa 8.400 mehr als 2018.

## Ausgangspunkt für die Etablierung des Netzwerks aus Forschungsdatenzentren<sup>2</sup>

Auf Initiative der Wissenschaft wurde 1999 die „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI) gegründet.<sup>3</sup> Dieses Gremium legte im März 2001 ein umfassendes Gutachten vor. Eine der zentralen Empfehlungen war die Einrichtung von FDZ bei den Statistischen Ämtern des Bundes und der Länder, der Deutschen Rente Bund und der

1 Eine alphabetische Auflistung der FDZ sowie Angaben zum Datenangebot der verschiedenen FDZ (Kategorien: Sozial, Wirtschaft, Bildung, Gesundheit, Psychologie, Qualitativ, Weitere) findet sich in Anhang B.

2 In Anhang A wird die Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD chronologisch übersichtlich dargestellt.

3 Zapf, Wolfgang et al. (1996): Memorandum zur Verbesserung der Zugangsmöglichkeiten zu Mikrodaten der amtlichen Statistik. ZUMA-Nachrichten 39, 172-175; Hauser, Richard; Gert G. Wagner und Klaus F. Zimmermann (1998): Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter wirtschafts- und sozialpolitischer Beratung. Allgemeines Statistisches Archiv 82(3), 369-379.

Bundesagentur für Arbeit, damit die dort vorhandenen Forschungsdaten professionell archiviert und soweit möglich etwa für die Replikation oder für die Beantwortung neuer Fragestellungen genutzt werden können. Im selben Jahr wurde der Gründungsausschuss des RatSWD konstituiert. Damit war der Grundstein für das dynamische Netzwerk rund um den RatSWD gelegt.

## Weiterentwicklung und Sicherung der Forschungsdateninfrastruktur

Seit 2001 entwickelte sich die Forschungsdateninfrastruktur stetig weiter. In diesem Zusammenhang gründeten sich auch neue FDZ, wobei die Motive zur Gründung eines FDZ dabei sehr heterogen waren: Die einen folgten externen Empfehlungen, wie solchen der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik, des Wissenschaftsrates oder wissenschaftlicher Beratungskreise. Andere erhielten aus der eigenen Einrichtung den Auftrag zur Gründung eines FDZ zum Zweck der Forschungsförderung. Allen gemein war jedoch das Bestreben, die Forschungsdateninfrastruktur in Deutschland durch die Bereitstellung von Forschungsdaten für die Wissenschaft zu erweitern und zu stärken.

Im Jahr 2004 wurde der RatSWD als strategisches Gremium gegründet. Um den operativen Dialog zwischen den FDZ zu fördern, wurde im Jahr 2009 der „Ständige Ausschuss Forschungsdateninfrastruktur“ (FDI Ausschuss) eingerichtet. Hauptaufgaben dieses Gremiums sind die kontinuierliche Sicherung und nachhaltige Verbesserung der Forschungsdateninfrastruktur: Dies beinhaltet den Ausbau der Qualität und Quantität des Datenangebots ebenso wie die Weiterentwicklung und Vereinfachung des Datenzugangs für die Forschungscommunity.

**79.878**

Downloads  
von frei verfügbaren  
Datensätzen

**2.359**

Publikationen  
die auf angebotenen  
Datensätzen basieren

### Infokasten 1:

#### Aufgaben und Struktur des RatSWD

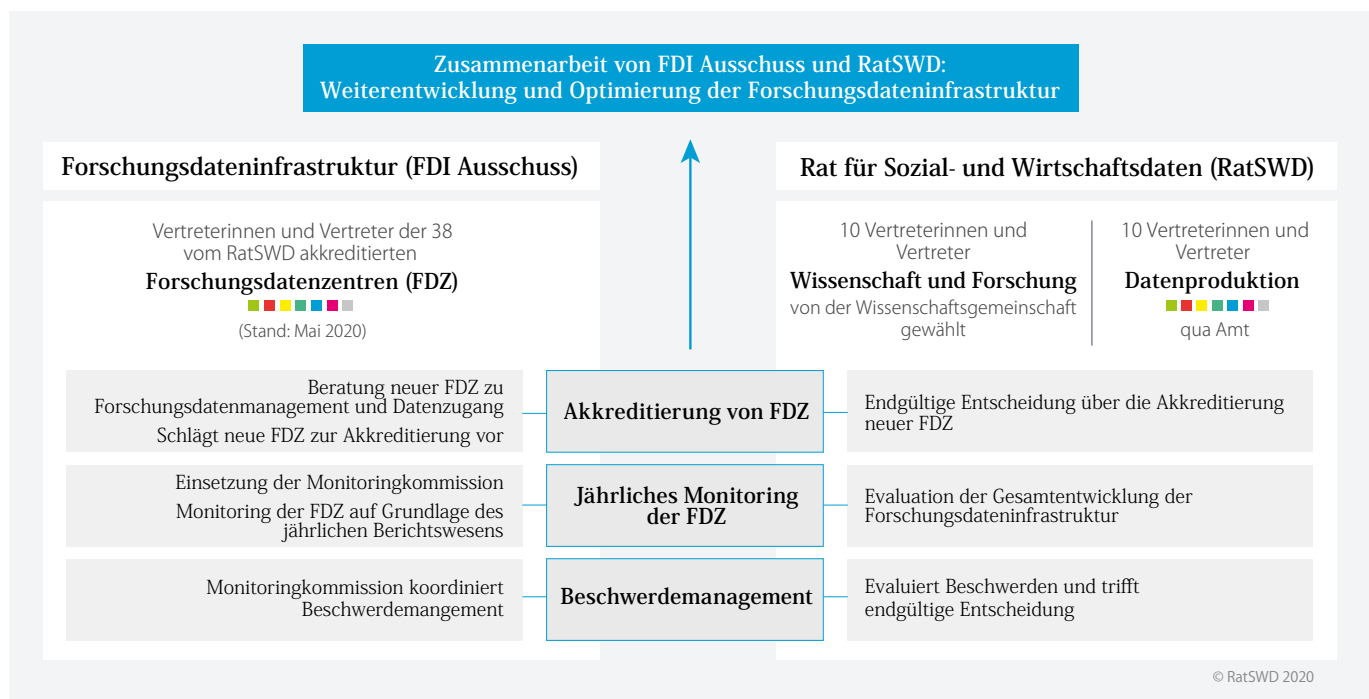
Der Rat für Sozial- und Wirtschaftsdaten (RatSWD) berät seit 2004 die Bundesregierung und die Regierungen der Länder in Fragen der Erweiterung und Verbesserung der Forschungsdateninfrastruktur für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften. Er ist ein institutionalisiertes Forum des Austausches und des Dialoges zwischen Wissenschaft und Datenproduktion über die Verbesserung des Zugangs zu qualitativ hochwertigen und wissenschaftlich interessanten Daten. Diese stammen aus den Tätigkeitsbereichen von staatlichen, wissenschaftsgetragenen und privatwirtschaftlichen Akteuren. Im RatSWD arbeiten (Stand der 7. Berufenungsperiode ab 2020) zehn durch Wahl legitimierte Vertreterinnen und Vertreter der sozial-, verhaltens- und wirtschaftswissenschaftlichen Fachdisziplinen sowie zehn Vertreterinnen und Vertreter der wichtigsten Datenproduzierenden zusammen. Mit seiner paritätischen Besetzung ist der RatSWD ein fachkundiges und gleichzeitig breit aufgestelltes Gremium. Es spielt eine tragende Rolle bei der Weiterentwicklung der Forschungsinfrastrukturen für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften und setzt sich für eine wissenschaftsfreundliche Gestaltung der rechtlichen und politischen Rahmenbedingungen ein.

Gleichzeitig koordiniert der RatSWD nunmehr 38 akkreditierte Datenzentren, welche im Ständigen Ausschuss Forschungsdateninfrastruktur (FDI Ausschuss) zusammenarbeiten. Er hat eine Infrastruktur geschaffen, die der Wissenschaft einen flexiblen und reichhaltigen Datenzugang ermöglicht.

## Akkreditierungsprozess und Qualitätssicherung

Zur Qualitätssicherung der Forschungsdateninfrastruktur hat der RatSWD im Jahr 2010 Mindeststandards und Kriterien für die Akkreditierung erarbeitet.<sup>4</sup> Seit 2015 werden diese laufend an technische Neuerungen und methodische Erweiterungen des Netzwerkportfolios angepasst. Durch ein jährliches Monitoring werden die Standards überprüft und die Qualität des Angebots gesichert. Die Sichtung der Akkreditierungsunterlagen neuer FDZ sowie das jährliche Monitoring werden von einer „Monitoringkommission“ realisiert, die aus dem FDI Ausschuss gewählt wird.<sup>5</sup> Über eine Akkreditierung entscheidet abschließend der RatSWD. Zusätzlich nehmen die FDZ an einem Beschwerdemanagement teil, das im Falle von nicht bilateral zu lösenden Fragen in der Datenbereitstellung bei der Geschäftsstelle des RatSWD administriert und von der Monitoringkommission verantwortet wird (vgl. Kapitel 8).

Abb. 1: Zusammenarbeit von FDI Ausschuss und RatSWD



<sup>4</sup> RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017): Qualitätssicherung der vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ). RatSWD Output 8 (5). doi: [10.17620/02671.4](https://doi.org/10.17620/02671.4)

<sup>5</sup> Hintergrundinformationen und die Mitglieder der Monitoringkommission sind im Anhang C zu finden.

Zusammengefasst werden die Ergebnisse des Monitorings in einem internen und dem hier vorliegenden öffentlich zugänglichen jährlichen Tätigkeitsbericht der vom RatSWD akkreditierten FDZ.<sup>6</sup> Die Akkreditierung durch den RatSWD ist für alle FDZ ein Qualitätssiegel, denn sie setzt die Einhaltung von Pflichtkriterien voraus: So muss zumindest ein Datenzugangsweg existieren, Datensätze müssen hinreichend dokumentiert sein, und es muss ein Konzept zur längerfristigen Verfügbarmachung der Daten vorgelegt werden. Ausschlaggebend für die Akkreditierung sind außerdem Informationen, etwa über vorhandene Arbeitshilfen, Qualitätssicherung und Angaben zur Weiterentwicklung der Infrastruktur und zur Sicherstellung des Datenschutzes.<sup>7</sup> Die Akkreditierung zahlt sich für die FDZ in verschiedener Hinsicht aus: Die FDZ werden infolge der Akkreditierung in ihrer Arbeit unterstützt, werden für die Weiterentwicklung der eigenen Infrastrukturen über Best Practice-Lösungen informiert und partizipieren am Erfahrungs- und Wissensaustausch mit den weiteren FDZ.

## Innovationspotential für die Forschung heben

Insbesondere aus dieser engen Zusammenarbeit entsteht erhebliches Innovationspotential für die Forschung. Die Bundesregierung und die Regierungen der Länder haben in der Gemeinsamen Wissenschaftskonferenz (GWK) 2018 auf Empfehlung des Rates für Informationsinfrastrukturen (RfII) die Einrichtung einer Nationalen Forschungsdateninfrastruktur (NFDI) im Rahmen einer Bund-Länder Vereinbarung beschlossen. Die NFDI soll nicht primär eine technische Infrastruktur sein, sondern einen Daten-Kulturwandel in der Wissenschaft herbeiführen und vernetzen, der das Dokumentieren und Teilen von Forschungsdaten als integralen Bestandteil der guten wissenschaftlichen Praxis betrachtet. Der RatSWD hat sich am durch die Deutsche Forschungsgesellschaft (DFG) koordinierten Ausschreibungsprozess 2019 mit der Konzeptualisierung eines Konsortiums für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (KonsortSWD) erfolgreich eingebracht. Der Prozess band den RatSWD und den FDI Ausschuss sowie zentrale in der Forschungsdateninfrastruktur etablierte Partner ein. Dadurch entwickeln sich künftig Services für die FDZ und vor allem für die Forschungs-Communities aus mehr als 15 sozial-, verhaltens- und wirtschaftswissenschaftlichen Fachgesellschaften.

<sup>6</sup> Die Tätigkeitsberichte seit dem Jahr 2015 sind abrufbar unter <https://www.ratswd.de/publikationen/taetigkeitsberichte>.

<sup>7</sup> Vgl. zu den Akkreditierungskriterien RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017): Qualitätssicherung der vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ). RatSWD Output 8 (5), 7. Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). doi: [10.17620/02671.4](https://doi.org/10.17620/02671.4).

## Neue Akkreditierungen

Im Jahr 2019 blieb der Kreis der 34 akkreditierten FDZ stabil, außerdem beantragten weitere FDZ eine Akkreditierung. Aufgrund der Vorarbeiten im Jahr 2019 konnten damit bis zum Sommer 2020 vier weitere FDZ vorläufig akkreditiert werden. Die neuen FDZ erweitern den Kreis der akkreditierten FDZ maßgeblich. So werden medizinisch-epidemiologische (GePaRD), finanzwirtschaftliche (SAFE FDZ) und raumbezogene (IÖR-Monitor) Daten, sowie Daten zu Migration und Integration (DeZIM.fdz) für die sozial-, verhaltens- und wirtschaftswissenschaftliche Forschung zugänglich.



### Pharmakoepidemiologische Forschungsdatenbank (GePaRD)

Das FDZ GePaRD (German Pharmacoepidemiological Research Database) ist eine pharmakoepidemiologische Forschungsdatenbank basierend auf Daten gesetzlicher Krankenversicherungen in Deutschland, welche seit 2004 am Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS aufgebaut und kontinuierlich gepflegt wird. Mit GePaRD können u. a. Forschungsfragen zum Gebrauch und zur Sicherheit von Arzneimitteln und Impfstoffen in der Routineversorgung untersucht werden.

GePaRD enthält Abrechnungsdaten von vier gesetzlichen Krankenversicherungen und umfasst Informationen von derzeit ungefähr 25 Millionen Personen, die seit 2004 oder später bei einer der teilnehmenden Krankenversicherungen versichert waren. Neben demographischen Angaben enthält GePaRD Informationen zu erstattungsfähigen Arzneimittelverordnungen sowie zu ambulanten und stationären Leistungen und Diagnosen. Pro Datenjahr stehen Informationen zu ungefähr 20 % der Allgemeinbevölkerung zur Verfügung und es sind alle geographischen Regionen Deutschlands vertreten. Der gesamte Prozess von der Datenlieferung über die Datenaufbereitung bis zur Verfügbarkeit der Daten für Auswertungen dauert meist 24 Monate, d. h. beispielsweise, dass Daten aus dem Jahr 2018 frühestens Ende 2020 verwendet werden können.

Das BIPS ist nicht Eigentümer der Daten und darf daher nicht entscheiden, für welche konkreten Projekte die Daten verwendet werden können. Die Genehmigung von Projekten basiert auf der Bewilligung des Vorhabens durch die Krankenversicherungen und die jeweils zuständigen Behörden. Entscheidend für die Nutzung der Daten gemäß § 75 SGB X ist, dass das öffentliche Interesse das schutzwürdige Interesse der betroffenen Personen erheblich überwiegt. Der Prozess der Genehmigung durch die Krankenversicherungen und die zuständigen Behörden dauert in der Regel mindestens drei Monate.

### Forschungsdatenzentrum des Leibniz-Institut für Finanzmarktforschung SAFE

Die wirtschaftswissenschaftliche Forschung im deutschen und europäischen Kontext wird dadurch erschwert, dass einerseits wenig paneuropäische Datensätze existieren und andererseits in den Datenbeständen bisher keine einheitliche Identifikation von z. B. Firmen möglich ist. Deshalb weichen Forscher oftmals auf US-amerikanische Daten aus, in denen die Integration verschiedener Datenbanken schon weiter fortgeschritten ist. Das hat aber zur Folge, dass empirische Analysen oftmals nur mithilfe nicht-europäischer Datensätze durchgeführt werden, sodass die daraus abgeleiteten Politikempfehlungen aufgrund der institutionellen Unterschiede nicht oder nur eingeschränkt auf den europäischen Raum übertragbar sind.

Vor diesem Hintergrund greift das SAFE Forschungsdatenzentrum nicht nur auf die üblichen internationalen Datenquellen zurück, sondern erhebt auch neue europäische Datensätze, führt vorhandene zusammen und bereitet sie auf. Ziel ist es, die fünf zentralen Forschungsbereiche von SAFE auf eine gemeinsame europäische Datenbasis zu stellen.



### Monitor der Siedlungs- und Freiraumentwicklung (IÖR-Monitor)

Der IÖR-Monitor ist eine Forschungsdateninfrastruktur des Leibniz-Instituts für ökologische Raumentwicklung (IÖR). Er stellt Daten und Informationen zur Flächennutzungsstruktur und deren Entwicklung sowie zur Landschaftsqualität für die Bundesrepublik Deutschland bereit.

Die flächenbezogenen Daten stehen hochaufgelöst auf geographischen Gitterzellen sowie unterschiedlichen administrativen Ebenen mit jährlicher Aktualität bereit. Die Datensätze basieren auf Forschungsergebnissen des IÖR. Kennblätter informieren als Metadaten über Datengrundlagen und Berechnungsmethodik. Grundlage für die Berechnung sind insbesondere geotopographische Daten (ATKIS Basis-DLM), Daten zur Landbedeckung (LBM-DE), amtliche Hausumringe (HU-DE) und Hauskoordinaten (HK-DE) sowie weitere Geofach- und Statistikdaten.

Die Datennutzung wird über einen Kartenviewer mit umfangreichen Interaktionsmöglichkeiten, einen Datenexport und über Geodienste angeboten. Die Daten können von interessierten Wissenschaftlerinnen und Wissenschaftlern für bestimmte räumliche Ausdehnungen und Zeiträume bedarfsgerecht bezogen werden. Ein Überblick über die mehr als 80 Datensätze samt Berechnungsmethoden, räumlichem und zeitlichem Bezug und den entsprechenden Exportfunktionen mittels Geodiensten ist unter [www.ioer-monitor.de/indikatoren](http://www.ioer-monitor.de/indikatoren) möglich.

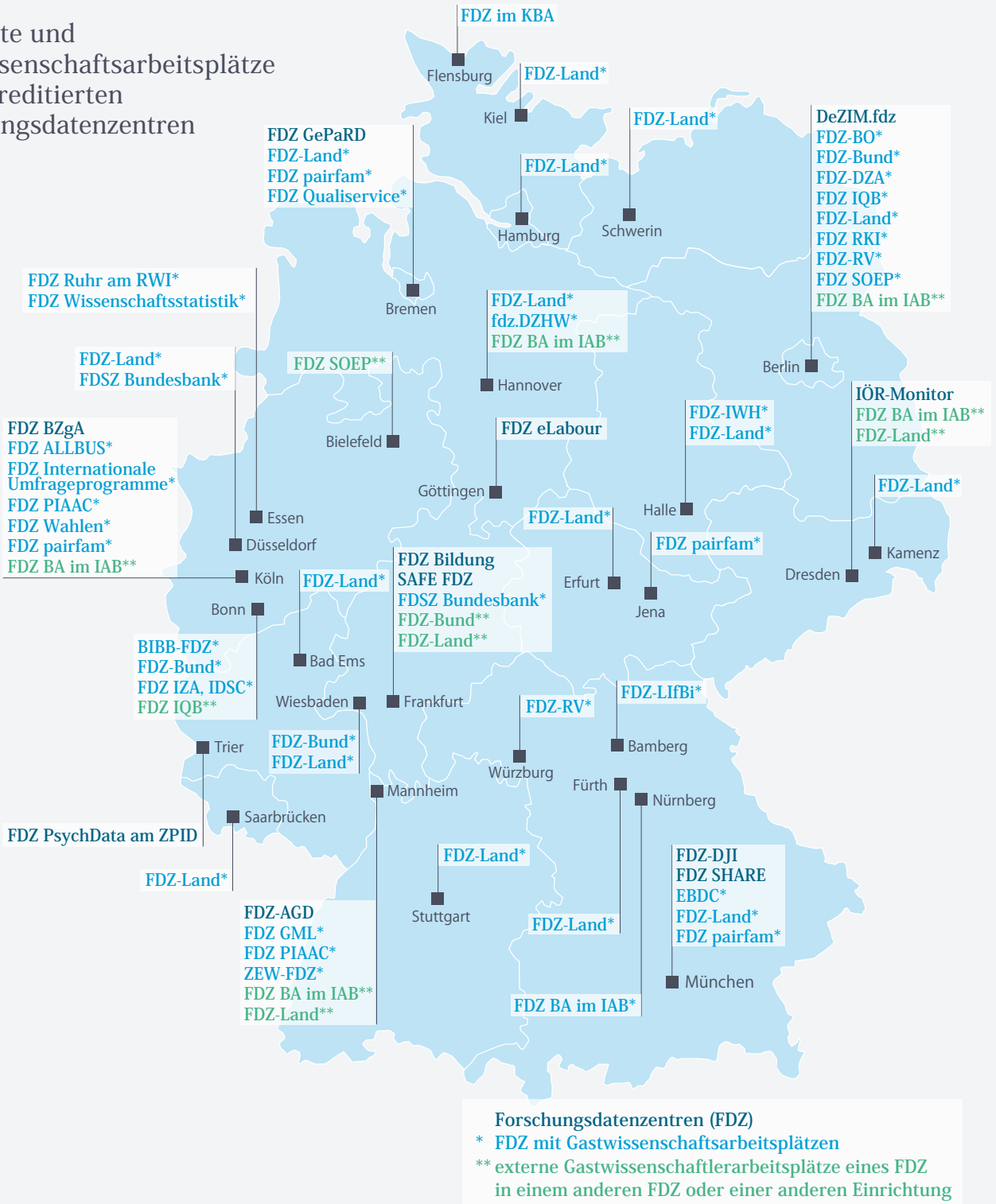


### Forschungsdatenzentrum Deutsches Zentrum für Integration und Migration (DeZIM)

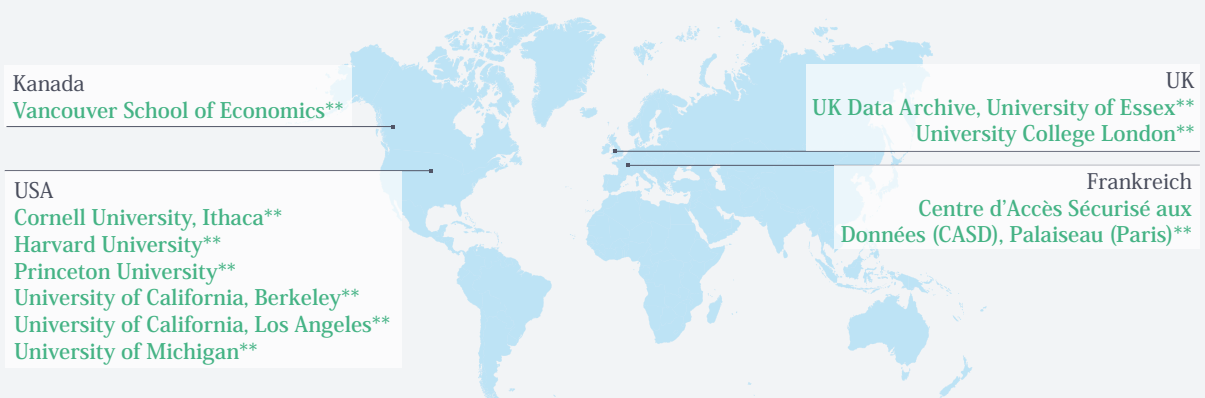
Das Deutsche Zentrum für Integrations- und Migrationsforschung (DeZIM) ist eine bundesweite Initiative von Wissenschaft und Politik zur Stärkung einer exzellenten und international sichtbaren Integrations- und Migrationsforschung in Deutschland. Das Forschungsdatenzentrum DeZIM.fdz bietet interessierten Wissenschaftlerinnen und Wissenschaftlern die Möglichkeit, auf Forschungsdaten zuzugreifen, die im Zuge der am DeZIM-Institut sowie an den Instituten der DeZIM-Gemeinschaft durchgeführten Projekte erhoben wurden. Neben dem Zugang zu den Daten bietet das DeZIM.fdz eine umfangreiche Informationsdatenbank an, die eine Recherche zu – intern wie auch extern archivierten – Studien ermöglicht, die sich mit den Themenbereichen Migration und Integration befassen. Weiterhin organisiert das DeZIM.fdz einen umfangreichen Support zu den einzelnen Datensätzen in seinem Datenangebot sowie zu verschiedenen methodischen Schwerpunktthemen.



# Standorte und Gastwissenschaftsarbetsplätze der akkreditierten Forschungsdatenzentren in 2020



## Internationale Standorte (FDZ BA im IAB)





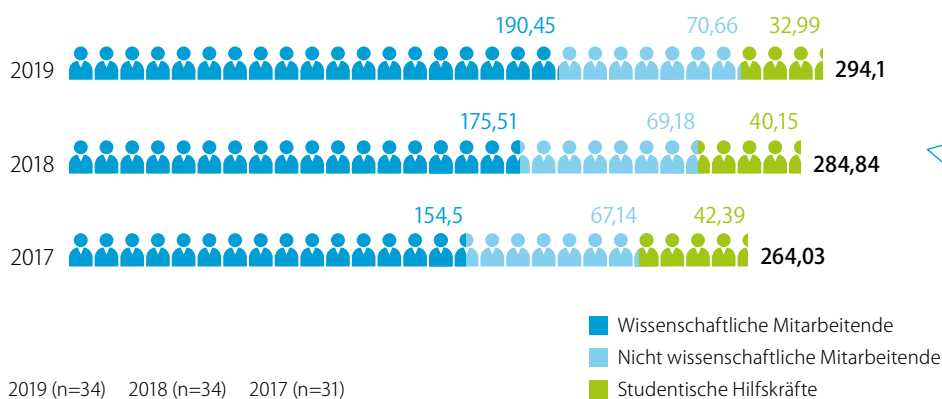
## 2 Struktur der Forschungsdatenzentren (FDZ)

Die in den folgenden Kapiteln zusammengefassten Informationen entstammen dem jährlichen Monitoring, das im Austausch mit allen Forschungsdatenzentren (FDZ) (weiter-)entwickelt wurde und an dem sich entsprechend auch alle vom RatSWD akkreditierten FDZ beteiligen. Im Berichtsjahr 2019 nahmen 34 FDZ am Monitoring teil.

### Personal

Die FDZ haben ihre personelle Ausstattung in den vergangenen Jahren kontinuierlich erweitert. Zum 31.12.2019 beschäftigten die 34 FDZ Personal in der Stärke von 294,10 Vollzeitäquivalenten (VZÄ) (siehe Abb. 3). Es ist also ein Personalzuwachs innerhalb der FDZ zu verzeichnen, der sich wie schon in den Jahren zuvor vor allem aus der Zunahme beim wissenschaftlichen Personal speist. Die Zahl der studentischen Hilfskräfte hingegen ist weiter rückläufig (vgl. Abb. 3). Ca. ein Drittel der FDZ arbeitet gänzlich ohne diese Statusgruppe.

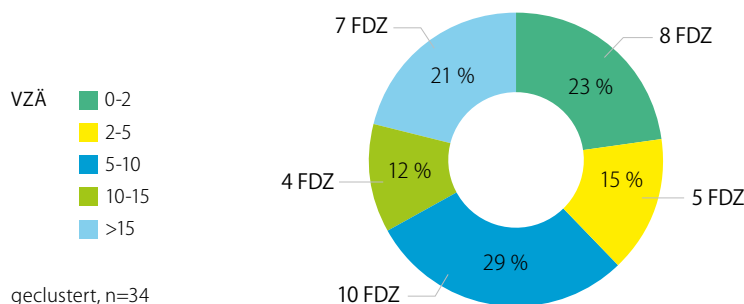
Abb. 3: Personal und Verteilung des Personalbestand in Vollzeitäquivalenten (VZÄ)



Über die letzten drei Jahre stieg der Personalbestand kontinuierlich

Die Anzahl der Mitarbeitenden in den FDZ variiert stark. Im Mittel verfügt ein FDZ über 8,9 Vollzeitäquivalente, tatsächlich reicht die Spannweite von weniger als einem VZÄ bis zu 36 VZÄ. Knapp zwei Drittel der FDZ haben maximal zehn Mitarbeitende in VZÄ, aber immerhin jedes fünfte FDZ hat mehr als 15 (vgl. Abb. 4).

Abb. 4: Personalbestand in Vollzeitäquivalenten (VZÄ)



Die Personalausstattung in den FDZ variiert stark, liegt aber bei den meisten FDZ bei fünf oder mehr VZÄ

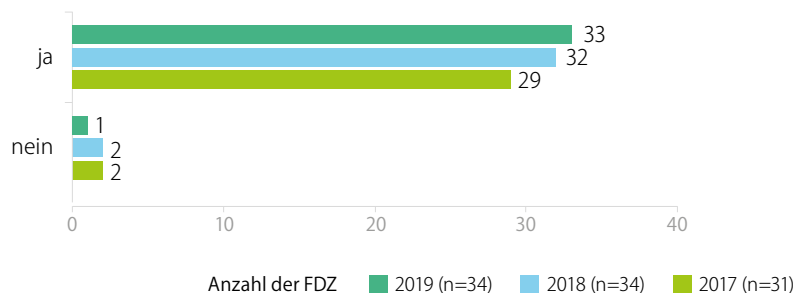


Da im Berichtszeitraum 2019 dieselben FDZ wie 2018 befragt wurden, ist der Personalaufwuchs nicht auf neu hinzugekommene FDZ zurückzuführen. Er kann vielmehr als Indiz gewertet werden, dass der Wandel in der Datenkultur, der vor einiger Zeit eingesetzt hat, sich auch in der Bereitschaft der Institutionen oder Mittelgeber niederschlägt, Personal für diese Aufgaben zu finanzieren. Die steigenden Nutzendenzahlen (siehe Kapitel 4) machen diesen Anstieg auch notwendig. Um die Datennutzenden kompetent und qualifiziert beraten zu können, ist es von großem Vorteil, wenn das FDZ-Personal selbst wissenschaftlich forscht. Dies ist in allen FDZ außer einem der Fall, wie Abb. 5 zeigt. Die Forschungstätigkeit in den FDZ bezieht sich sowohl auf inhaltliche Fragestellungen als auch auf methodische oder technologisch-funktionelle Weiterentwicklungen im Bereich Dateninfrastruktur und Datenservice.

**Abb. 5: Selbstständige Forschung von wissenschaftlichem Personal**

Gibt es wissenschaftlich ausgebildetes Personal am FDZ, das selbstständig forscht?

In nahezu allen FDZ forschen wissenschaftlich ausgebildete Mitarbeitende auch selbstständig



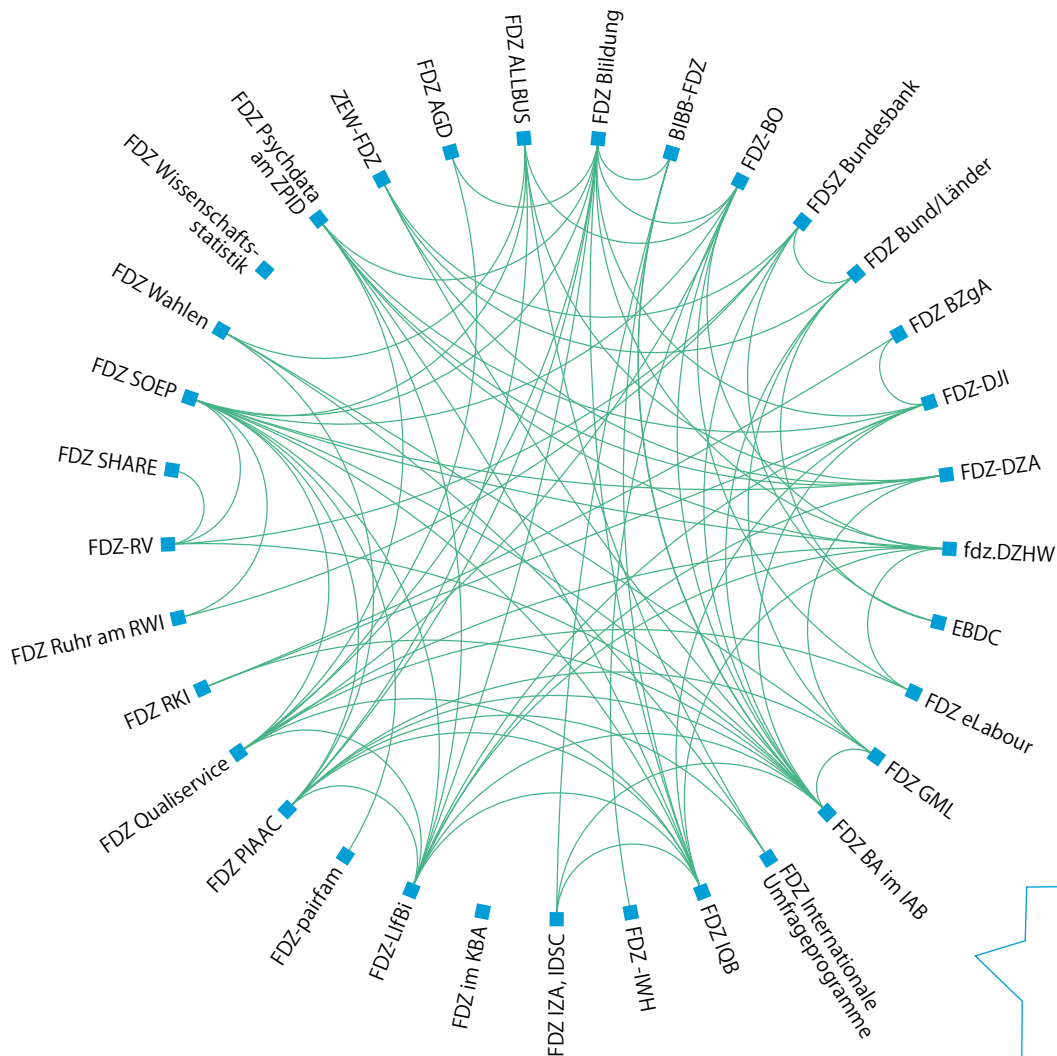
© RatSWD 2020

Zum Umfang dieser Tätigkeit: Feste Arbeitszeitanteile für eigene Forschung des wissenschaftlichen Personals gibt es in 16 FDZ. Die Anteile betragen im Durchschnitt ein knappes Drittel der Arbeitszeit, mit einer breiten Streuung in den einzelnen Datenzentren. In sechs FDZ wird bis zu einem Viertel der Arbeitszeit dafür aufgewendet, in zehn FDZ liegt der Anteil darüber.

### Kooperationen und Forschungsaktivitäten

Die Anzahl der FDZ, die im Inland Kooperationsbeziehungen in einer institutionalisierten Form pflegen, hat sich gegenüber dem Vorjahr erhöht. 25 FDZ geben Kooperationen an und im Gegenzug hat sich die Zahl derjenigen, die derzeit keine Kooperationen haben und das momentan auch nicht planen, von elf FDZ auf acht reduziert (zu den internationalen Forschungsk Kooperationen siehe Kapitel 6, S. 38). Über die Mitarbeit im FDI-Ausschuss, aber auch darüber hinaus, kooperieren alle beim RatSWD akkreditierten FDZ zudem untereinander, wie Abb. 6 zeigt.

Abb. 6: Kooperationen zwischen den akkreditierten FDZ im Jahr 2019



Der Austausch im FDI Ausschuss und darüber hinausgehende Kooperationen ergeben ein vielschichtiges interdisziplinäres Netzwerk an FDZ

Diese Kooperationen können sowohl inhaltlich motiviert sein und FDZ bündeln, die ihre Dateninfrastruktur auf ein Thema fokussieren (bspw. der Verbund FDB, welcher Forschungsdaten für die empirische Bildungsforschung bereitstellt) oder die eine methodische Ausrichtung haben. Ein Beispiel hierfür ist VQualidat, ein Verbund, in welchem sich FDZ zusammengeschlossen haben, die qualitativ erhobene Forschungsdaten anbieten. Angestrebt wird außerdem der Ausbau des Konzeptes der externen Gastwissenschaftsarbeitsplätze (FDZ im FDZ). Hierbei handelt es sich um die Möglichkeit, an einem akkreditierten FDZ auch Zugang zu Daten eines anderen FDZ zu bekommen. Nicht zuletzt wird die Zusammenarbeit der FDZ untereinander weiter durch das Konsortium für Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (KonsortSWD) vertieft werden, welches im Rahmen der Nationalen Forschungsdateninfrastruktur (NFDI) gefördert wird. Das Netzwerk der im Betrieb von nutzungsorientierten Forschungsdateninfrastrukturen erfahrenen FDZ bildet das operative Rückgrat des KonsortSWD (vgl. zu KonsortSWD auch S. 11).

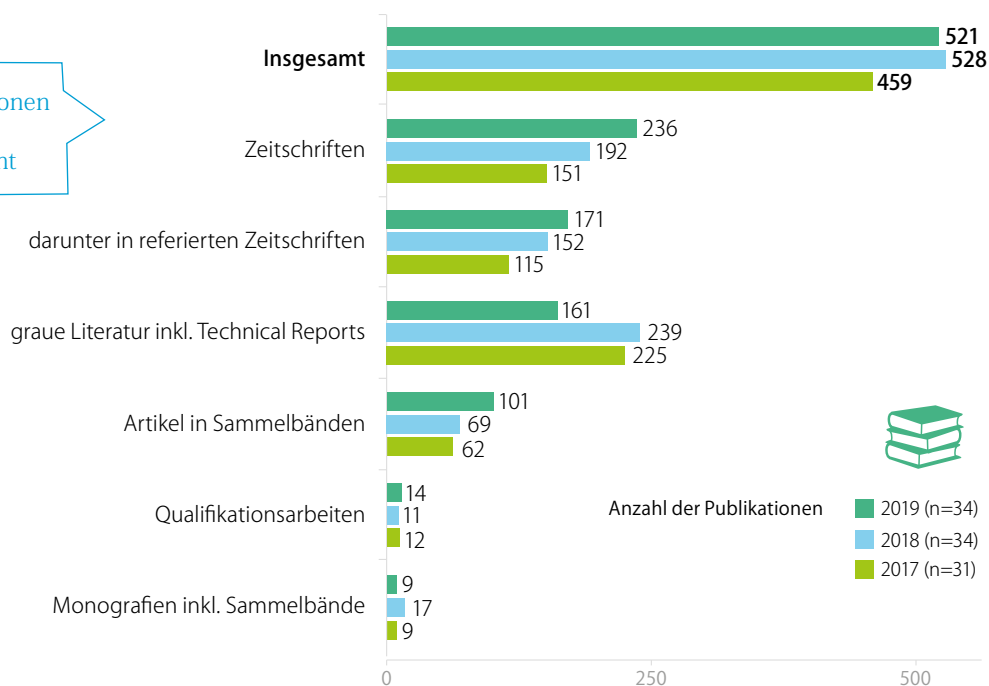
## Wissenschaftliche Publikationen

Tätigkeiten im Forschungsdatenzentrum werden in erster Linie mit der Bereitstellung von Daten für nichtkommerzielle wissenschaftliche Forschung assoziiert. Mit der Bereitstellung der Daten geht in der Regel auch ein Beratungsangebot zu den Potentialen der bereitgestellten Datensätze einher. Daher besteht die Mehrheit des FDZ-Personals aus wissenschaftlichen Mitarbeitenden und ein Teil ihrer Arbeitszeit wird eigener Forschungstätigkeit gewidmet. Diese Aktivitäten dienen letztlich dazu, die Daten-, Service- und Beratungstätigkeiten stets auf angemessenem wissenschaftlichem Niveau durchzuführen. Sie werden auch in wissenschaftlichen Publikationen dokumentiert. Insgesamt ist der Publikationsoutput in etwa auf Vorjahresniveau geblieben, wie Abb. 7 zeigt. Im Berichtszeitraum wurden 521 Publikationen verfasst.

**Abb. 7: Wissenschaftliche Publikationen des FDZ Personals**

Bitte geben Sie die Anzahl der wissenschaftlichen Publikationen Ihrer FDZ-Mitarbeitenden an, unabhängig von den verwendeten Daten und unabhängig davon, ob die Publikation im Rahmen der FDZ-Arbeitszeit verfasst wurde.

Ein Drittel aller Publikationen wurde in referierten Zeitschriften veröffentlicht



Deutlich gesteigert wurde im Jahr 2019 der Anteil von Publikationen in Zeitschriften. Vor allem ist es gelungen, noch mehr Artikel in referierten Zeitschriften zu veröffentlichen, was als Ausweis der hohen Qualität der Inhalte gewertet werden kann. Die Zahl der Qualifikationsarbeiten ist ebenfalls gestiegen. Dabei handelt es sich vor allem um höhere Qualifikationen (Promotion, Habilitation; in der Abbildung nicht extra ausgewiesen). Masterarbeiten zählen auch zu den Qualifikationsarbeiten, werden aber nicht von jedem FDZ reportiert, weshalb an dieser Stelle von einer Untererfassung ausgegangen werden kann. Das heißt aber nicht, dass im akademischen Qualifikationsverlauf „niedrigeren“ Abschlussarbeiten weniger Bedeutung beigemessen wird. Im Gegenteil, seit Juni 2020 erstattet das Statistische Bundesamt bei Qualifikationsarbeiten die bereits vergünstigten Kosten bei der Datennutzung für Nachwuchswissenschaftlerinnen und -wissenschaftler komplett. Somit kann der Datenzugang für Bachelor, Master- und Promotionsarbeiten im Statistischen Verbund kostenlos angeboten werden.

### 3 Archivierung und Qualitätssicherung

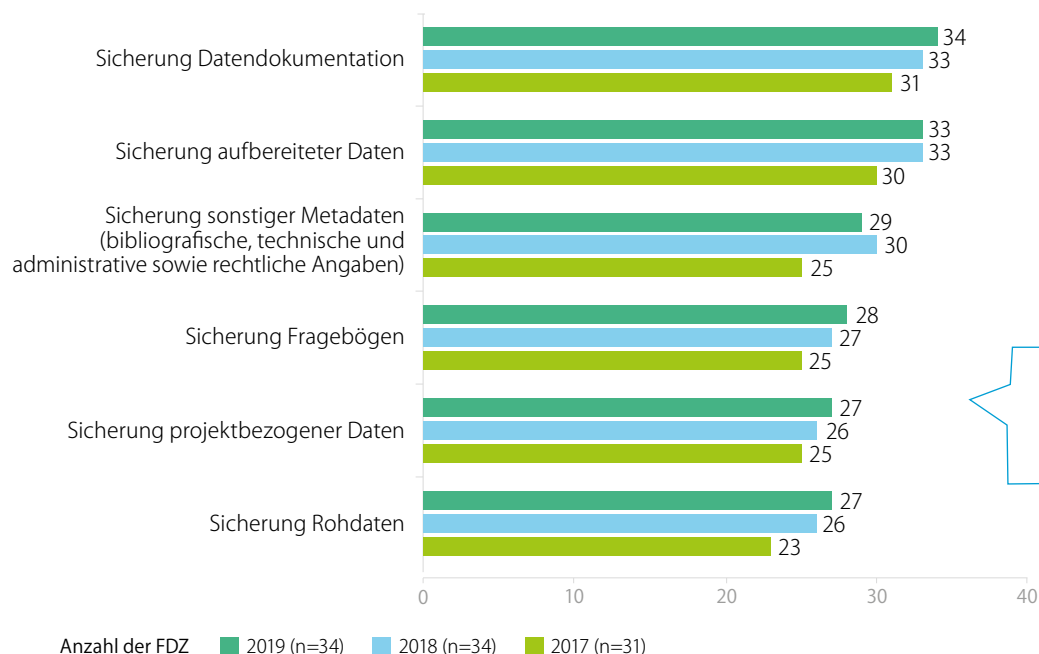
#### Konzepte für die langfristige Verfügbarmachung der Daten

Die langfristige Verfügbarkeit der für die Forschung verwendeten Daten ist Teil der guten wissenschaftlichen Praxis und deshalb in den Forschungseinrichtungen fest verankert. So sichern alle FDZ ihre Datendokumentationen und auch die aufbereiteten Daten (siehe Abb. 8, für ein FDZ ist dieser zweite Punkt nicht relevant). Die langfristige Sicherstellung der Daten nach ihrer Primärverwendung dient unter anderem dem Nachweis der geleisteten Forschungstätigkeit, der Replizierbarkeit der empirischen Forschungsergebnisse (bspw. bei Qualifizierungsarbeiten oder Publikationen in Zeitschriften) und auch der Nachnutzung durch Dritte.

Um dies zu gewährleisten, ist nicht nur die physische Speicherung der Daten, ggf. auch in unterschiedlichen Versionierungen notwendig, sondern auch die längerfristige Sicherstellung ihrer Lesbarkeit, Interpretierbarkeit und Nutzbarkeit. Die Datenumgebung kann sich über die Zeit ändern (z. B. durch neue Versionen bei der verwendeten Statistiksoftware, bei der eine Abwärtskompatibilität zu älteren Fassungen nicht immer gewährleistet ist). Eine wesentliche Herausforderung besteht deshalb darin, die Daten in Formaten vorzuhalten, die auch mit zukünftigen Technologien noch lesbar sind. Außerdem ist auch ein Schutz vor physischem Datenverlust sowie unbefugter bzw. ungewollter Manipulation notwendig.

**Abb. 8: Datensicherung in den FDZ**

Stellt Ihr FDZ sicher, dass die Daten längerfristig (entsprechend der Regeln guter wissenschaftlicher Praxis mindestens 10 Jahre nach der letzten Nutzung) und in gängigen Formaten verfügbar sind?  
(Mehrfachantworten möglich)



Die langfristige Datensicherung ist eine Kernaufgabe in den FDZ

Nicht alle FDZ geben an, dass die Sicherung von Rohdaten oder Fragebögen zu ihren Aufgaben gehört. Das heißt aber nicht, dass diese zwei Elemente nicht gesichert werden, sondern es bedeutet, dass diese Aufgaben an anderer Stelle erledigt werden.

## Archivierung

In 33 FDZ werden die Daten redundant auf mehreren Datenträgern gesichert und 17 dieser FDZ geben an, dass die Sicherung auch an unterschiedlichen Orten erfolgt. Ein FDZ nutzt die einfache Sicherung auf institutsinternen Servern. Bei manchen empirischen Studien können Updates von Daten oder Metadaten notwendig werden. In den FDZ wird damit unterschiedlich verfahren. Am häufigsten werden alle Versionen des Datensatzes archiviert und den Nutzenden zur Verfügung gestellt (in 18 FDZ). In elf FDZ steht dann jeweils die aktuelle Version zur Nutzung bereit und fünf FDZ verwalten Datensätze, bei denen keine Updates erforderlich sind.

Die Zahl der FDZ mit einer Zertifizierung für ihre Archivierungslösung hat sich zum Vorjahr nicht geändert und beträgt zehn FDZ. Mehrheitlich haben sie ein Zertifikat von „CoreTrustSeal“, einer community-basierten, gemeinnützigen Organisation. Ziel der Zertifizierungen ist es, den Forschenden eine Orientierungshilfe zur Auswahl der Repositorien für die Archivierung und Kuratierung der Forschungsdaten zu geben. Entsprechend geben diese Zertifizierungen Auskunft über die Einhaltung transparenter (und in den Zertifizierungskriterien festgelegter) Standards in den Repositorien. Zur Förderung der Zertifizierung in den FDZ und um die damit verbundenen Prozesse zu erleichtern, soll es in KonsortSWD ein gesondertes Arbeitspaket geben. Im Mittelpunkt steht hier neben der (Neu-)Zertifizierung von FDZ durch CoreTrustSeal auch die Akkreditierung durch den RatSWD. Das Arbeitspaket möchte also auch speziell für FDZ, die bisher nicht im Netzwerk um den RatSWD sind, ein Angebot machen und die Akkreditierungsvorbereitung unterstützen.

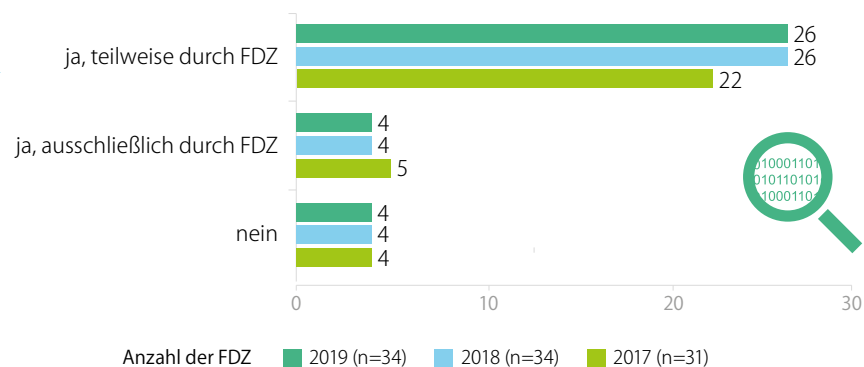
## Qualitätssicherung der Datensätze

Nicht nur die physische Sicherung der Daten für Nutzung und Nachnutzung gehört zu den Aufgaben der FDZ, sondern in den meisten Fällen auch die Mitarbeit bei der Datenprüfung. Wie die FDZ in diese Arbeitsschritte eingebunden sind, hängt u. a. damit zusammen, ob und wie das Datenzentrum mit einer weiteren (Träger)Institution verbunden ist. Da diese Einbindung über die Zeit üblicherweise stabil ist, hat sich im Berichtszeitraum nichts Wesentliches an den Aufgabenverteilungen geändert. Das zeigt auch Abb. 9. Wie schon im Vorjahr gibt es vier FDZ, denen die Datenprüfung allein obliegt und in knapp 80 % der Datenzentren erfolgt die Prüfung auf Qualität und Güte der Daten teilweise durch das FDZ. In vier Datenzentren werden keine Datenprüfungen vorgenommen.

**Abb. 9: Datenprüfung in den FDZ**

Gehört die Datenprüfung (auf Qualität und Güte der weitergegebenen Daten) zur Aufgabe Ihres FDZ?

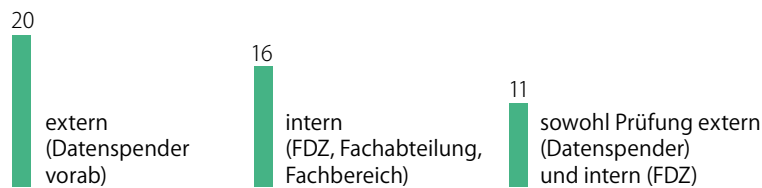
Die Mehrheit der FDZ ist in die Datenprüfung eingebunden



Von denjenigen FDZ, denen die Datenprüfung nicht allein obliegt, werden am häufigsten die Primärforschenden oder Datenproduzierenden (dies können auch Drittinstitute sein) als weitere Akteure bei der Qualitätssicherung benannt (siehe Abb. 10). In 18 FDZ gibt es ein Regelwerk für die Datenprüfung.

**Abb. 10: Von wem (noch) und wie wird die Datenprüfung durchgeführt?**

(Auswertung offene Frage, Mehrfachnennungen möglich)



Die Datenprüfung erfolgt extern und intern

■ Anzahl der FDZ in 2019 (n=30)

© RatSWD 2020

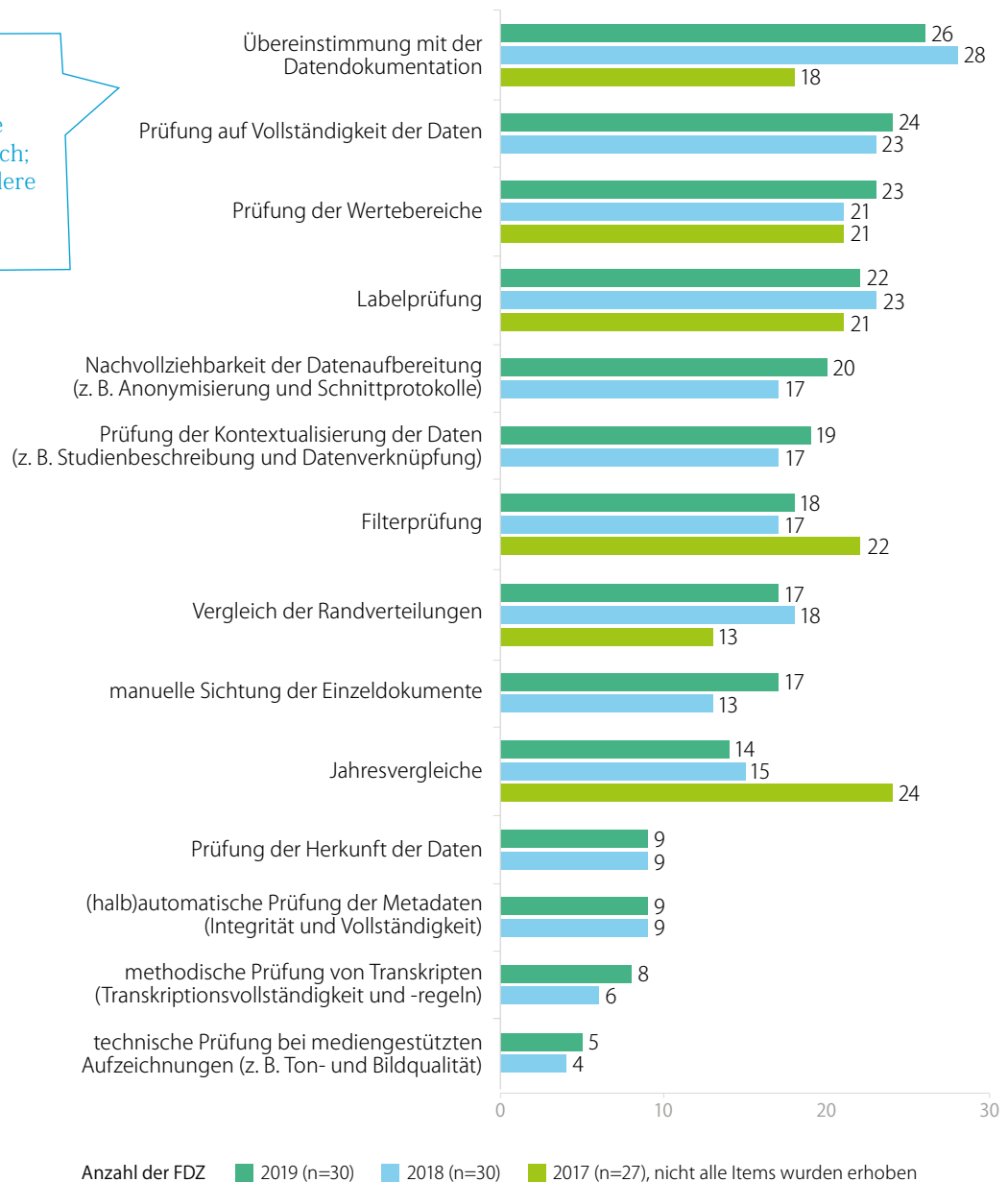
Am häufigsten prüfen die FDZ auf Übereinstimmung mit den Datendokumentationen, auf Vollständigkeit der Daten und die Korrektheit der Wertebereiche (siehe Abb. 11). Die Anzahl von FDZ, die angeben, die Herkunft der Daten zu kontrollieren oder auch einen automatischen Abgleich der Metadaten vorzunehmen, ist hingegen deutlich geringer. Ein Grund dafür kann sein, dass die Herkunft der Daten beispielsweise bereits vor der Übergabe an das FDZ geprüft wurde. Wenn bestimmte Datenprüfungen nicht im FDZ vorgenommen werden, heißt das nicht, dass sie nicht stattfinden, sondern an anderer Stelle im Lebenszyklus der Daten verankert sind. Die methodische Prüfung von Transkripten und die technische Prüfung mediengestützter Aufzeichnungen werden in wenigen FDZ durchgeführt. Das sind Arbeitsschritte, die in der qualitativen empirischen Forschung zum Einsatz kommen, worauf nur ein kleiner Teil der akkreditierten FDZ spezialisiert ist.

**Abb. 11: Art der Datenprüfung**

Welche Art der Datenprüfung wird in Ihrem FDZ durchgeführt?

(Mehrfachantworten möglich)

Die FDZ führen für die jeweiligen Datensätze passende Datenprüfungen durch; diese sind insbesondere abhängig von der Erhebungsmethode

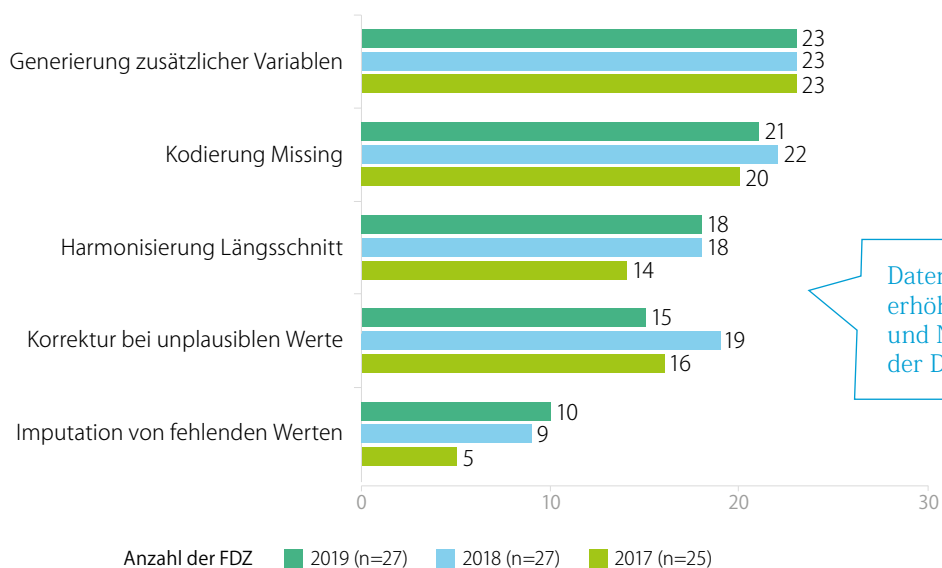


Die Datenprüfung ist ein wichtiger Schritt in der Qualitätssicherung der Daten. Ein weiterer Bestandteil ist die Korrektur von Datenfehlern einschließlich der Generierung von Daten, z. B. um die Nutzungsfreundlichkeit zu erhöhen. Hier ist die Aufgabenverteilung ähnlich wie bei der Datenprüfung: In zwei FDZ obliegt die Korrektur ausschließlich, in 25 Fällen teilweise den FDZ. In sieben FDZ gehört die Datenkorrektur nicht zum Tätigkeitsspektrum.

Die häufigste Aufgabe in diesem Bereich ist die Generierung zusätzlicher Variablen, 23 FDZ geben dies an. Aber auch die Kennzeichnung fehlender Werte durch entsprechende Kodierungen wird häufig vorgenommen (21 FDZ). Die Längsschnittharmonisierung gehört bei 18 FDZ zu den qualitätssichernden Korrekturen (vgl. Abb. 12). In 15 der 27 FDZ, die Datenkorrekturen durchführen, gibt es dafür ein festes Regelwerk. Für die Datennutzenden sind die vorgenommenen Korrekturen nicht immer auch auf der Einzelfallebene erkennbar, z. B. wenn die Korrekturen aus Gründen des Datenschutzes oder der Anonymisierung erfolgten.

**Abb. 12: Art der Datenkorrektur in den FDZ**

Welche Art der Generierung oder Datenkorrektur wird in Ihrem FDZ durchgeführt?  
(Mehrfachantworten möglich)



Datenkorrekturen erhöhen die Qualität und Nutzbarkeit der Daten



## 4 Datenangebot und Datennutzung

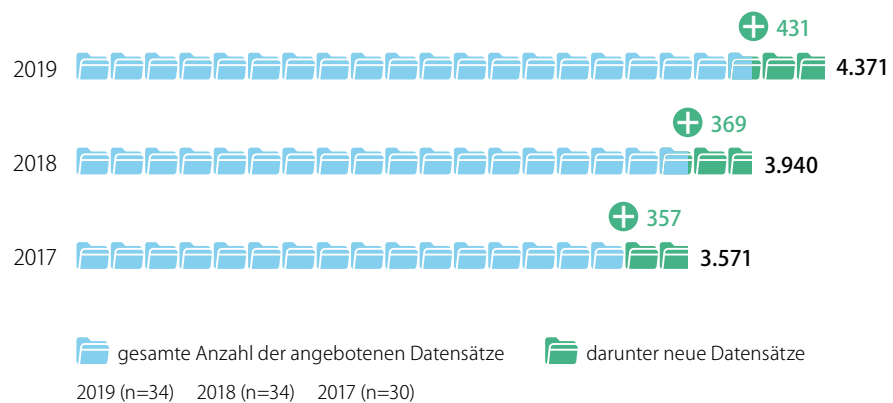
### Umfang des Datenangebotes

Der Zuwachs im Datenangebot der letzten Jahre setzt sich fort: Zum Stichtag 31.12.2019 haben die Forschungsdatenzentren (FDZ) 4.371 Datensätze angeboten. Damit sind im Berichtsjahr 2019 431 Datensätze neu hinzugekommen, die entweder einen neuen Digital Object Identifier (DOI) erhalten haben oder prinzipiell dazu geeignet gewesen wären (siehe Abb. 13). Hierbei variiert in den FDZ die Anzahl der neu angebotenen Datensätze, beginnend bei einer einstelligen Anzahl bis zu 151 neuen Datensätzen in einem FDZ. Ein Datensatz kann aus mehreren Einzelstudien bestehen, so dass die Anzahl bereitgestellter Studien wesentlich höher liegt.

**Abb. 13: Angebotene Datensätze in den FDZ**

Bitte geben Sie die im Berichtsjahr hinzugekommenen Erhebungen und Datensätze an.  
Hinweis: Als neu gelten Datensätze, die z. B. eine neue DOI erhalten haben oder prinzipiell dazu geeignet wären.

Der Anteil der angebotenen Datensätze stieg in den vergangenen Jahren an



© RatSWD 2020

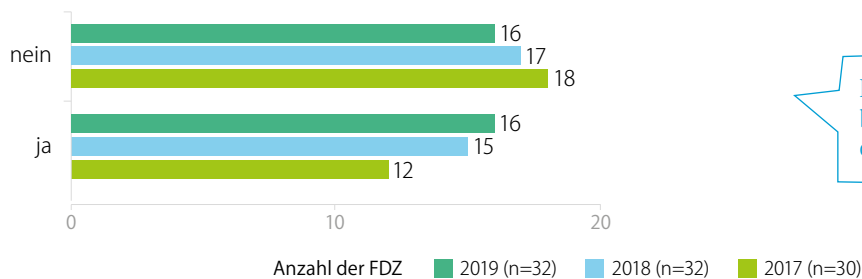
Für die angebotenen Datensätze vergeben 31 FDZ Persistente Identifikatoren (PID), um deren Auffindbarkeit und Zitierbarkeit dauerhaft zu gewährleisten. Alle FDZ, die noch keine PID, wie z. B. DOI, vergeben, planen aktuell die Einführung eines solchen Verfahrens.

### Zeitpunkt der Verfügbarmachung der Datensätze und Gebühren

Das zentrale Ziel der FDZ ist ein niedrigschwelliger und zeitnaher Datenzugang. Um die Chancengerechtigkeit zu wahren, sollten die Forschungsdaten allen qualifizierten Nutzenden zeitgleich zugänglich gemacht werden. Hohe Gebühren und lange Sperrfristen zur Datenweitergabe können diesen Zielen entgegenstehen. Gleichzeitig können aber auch gute Gründe vorliegen, warum ein zeitliches Embargo für die Datenweitergabe verhängt wird. So begründen die meisten FDZ Embargos mit Erstverwertungsmöglichkeiten und -rechten der Primärforschenden. Es werden damit etwa laufende Qualifikationsarbeiten durch Sperrfristen geschützt.

**Abb. 14: Sperrfristen für Datensätze in den FDZ**

Gibt es Sperrfristen für einige Datensätze?

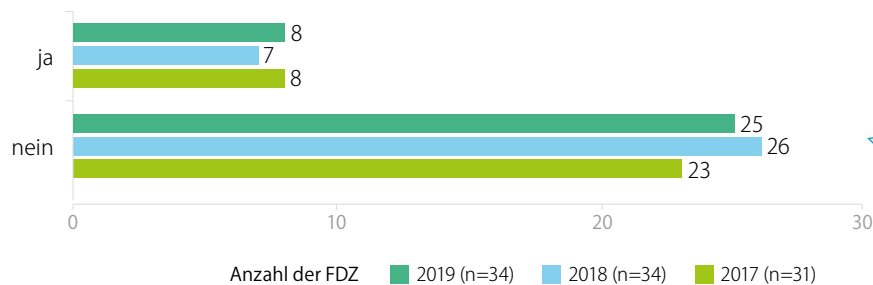


Die Hälfte der FDZ bieten Datensätze ohne Sperrfristen an

In der Hälfte der FDZ existieren keine Sperrfristen, d. h. die Daten werden unmittelbar nach der Aufnahme und Aufbereitung zur Verfügung gestellt (siehe Abb. 14). Ebenfalls 16 FDZ gaben 2019 an, dass sie zumindest einen Teil ihrer Datenbestände mit Sperrfristen versehen. Fünf FDZ nennen feste Sperrfristen von sechs Monaten bis maximal zwei Jahren. Bei den übrigen FDZ sind die Sperrfristen nicht grundsätzlich festgelegt, sondern abhängig von bestimmten Bedingungen, wie beispielsweise dem Laufzeitende von Forschungsprojekten oder Vorgaben von fördernden Institutionen bzw. den Datengebern selbst. Insgesamt ist über die letzten Jahre eine gewisse Dynamik in der Nutzung von Sperrfristen zu bemerken, die absolute Anzahl der FDZ, die Datensätze mit einem zeitlichen Embargo versehen, ist leicht gestiegen.

**Abb. 15: Gebühren der Datenbereitstellung in den FDZ**

Erheben Sie Gebühren bei der Datenbereitstellung für die wissenschaftliche Nutzung?



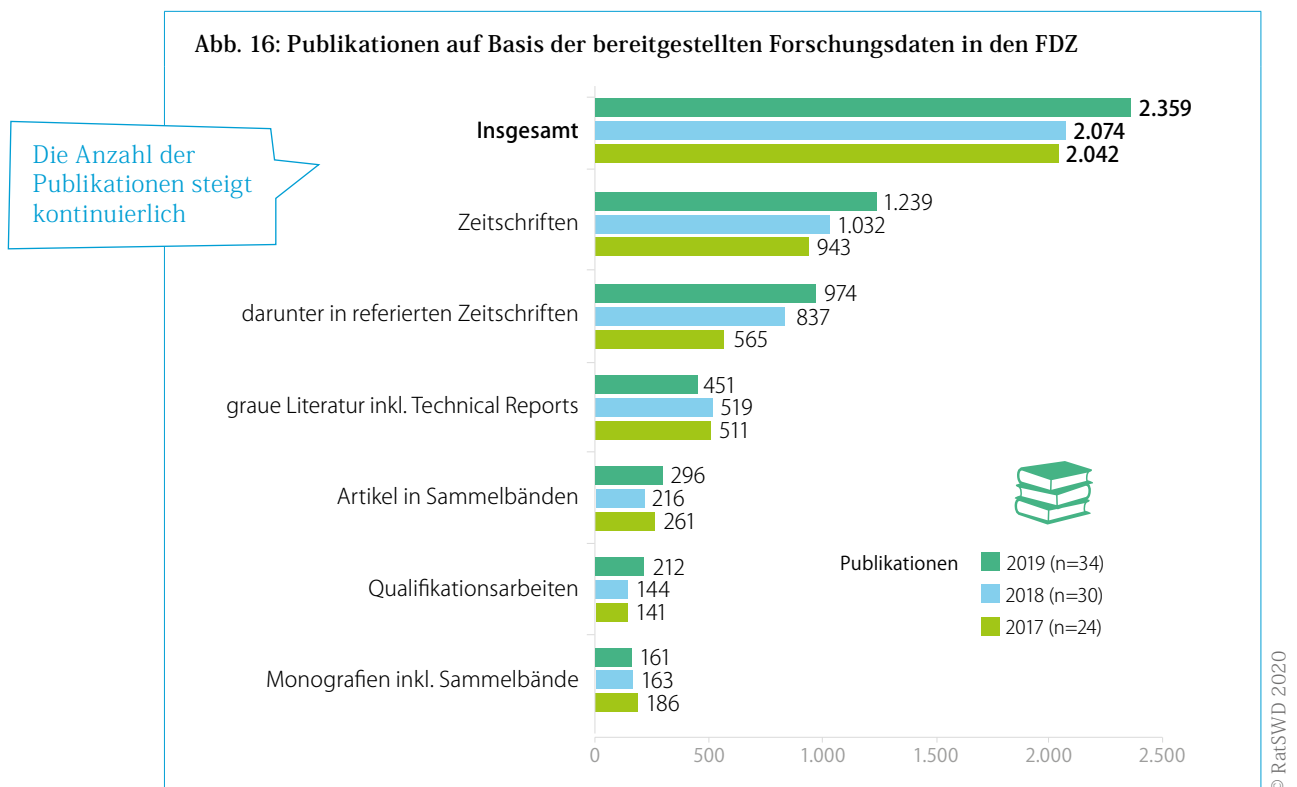
Die große Mehrheit der FDZ erhebt keine Gebühren bei der Datenbereitstellung

Neben den Sperrfristen kann eine Nutzungsgebühr ein Hemmnis für die Datennutzung sein. Ein Großteil der FDZ verzichtet völlig auf eine solche Gebühr: Von 34 FDZ gaben 25 für das Jahr 2019 an, dass sie keine Gebühren erhoben haben (siehe Abb. 15). Verglichen mit dem Vorjahr zeigt sich hier kaum Veränderung. Die genannten Gebühren der acht FDZ liegen zum größten Teil im zweistelligen bzw. unteren dreistelligen Euro-Bereich, häufig mit Reduktionen für die Durchführung von Promotionsarbeiten.<sup>8</sup> Wofür die Gebühren erhoben werden, differiert dabei zwischen den FDZ: Teilweise werden die Gebühren pro Datensatz oder Zugangsweg erhoben, teilweise wird die Datennutzung pro Erhebungsjahr berechnet. Auch der Aufwand für die Datenaufbereitung, beispielsweise für eine speziell notwendige Anonymisierung, oder Beratungsleistungen, sind teilweise kostenpflichtig. Gerade die niedrigen Beträge dürften allerdings nur wenig zur Kostendeckung des Aufwands in den FDZ beitragen. Sie dienen dann jedoch im Sinne einer Schutzgebühr auch dazu, die Ernsthaftigkeit der Datennutzung sicherzustellen und „Fake-Bestellungen“ zu vermeiden.

<sup>8</sup> Seit Juni 2020 erstattet das Statistische Bundesamt bei Qualifikationsarbeiten die bereits vergünstigten Kosten bei der Datennutzung für Nachwuchswissenschaftler und -wissenschaftlerinnen komplett. Somit kann der Datenzugang für Bachelor-, Master- und Promotionsarbeiten im Statistischen Verbund kostenlos angeboten werden.

## Forschungsooutput auf Basis der bereitgestellten Forschungsdaten

Der zentrale Indikator für die Datennutzung ist der Forschungsooutput. Die Anzahl der Publikationen ist ein wesentlicher Indikator für diese Messgröße. Dabei müssen qualitative Unterschiede zwischen den Publikationsformen unterschieden werden. So sind Artikel in referierten Zeitschriften als höherwertig zu betrachten als in Fachzeitschriften ohne Peer-Review-Verfahren. Artikel in Fachzeitschriften stellen auch im Jahr 2019 die häufigste von den Datennutzenden gemeldete Publikationsart dar. Der Anteil in referierten Fachzeitschriften ist besonders hoch und weist auch die höchsten Zuwächse aus.



Im Jahr 2019 meldeten die 34 FDZ insgesamt 2.359 Publikationen (siehe Abb. 16), die auf Basis der bereitgestellten Forschungsdaten erschienen sind. Dabei ist zu berücksichtigen, dass Forschende häufig die Rückmeldung oder Abgabe von Belegexemplaren an die FDZ versäumen. Obwohl die FDZ mit der breiten Nutzung persistenter Identifikatoren für Forschungsdaten die Möglichkeit zur Datenzitation gelegt haben, berücksichtigen nicht alle Autorinnen und Autoren diese Zitationsmöglichkeiten in ihren Publikationen. Da zudem nicht jedes FDZ die Ressourcen für eine eigene Recherche von Datensatzzitationen hat, ist von einer erheblichen Untererfassung auszugehen. Dazu trägt auch bei, dass mehrere Nutzende unter Umständen mehrere Datensätze aus verschiedenen FDZ verwenden.

Auch wenn die Ermittlung der Anzahl der auf der Grundlage der bereitgestellten Daten verfassten Publikationen also nicht unproblematisch ist, kann im langjährigen Vergleich insgesamt ein Anstieg der erfassten Publikationen verzeichnet werden. Wie bereits ausgeführt gilt dies vor allem für Zeitschriftenartikel. Aber auch die Anzahl der Qualifikationsarbeiten, die mit den Daten aus den FDZ erstellt werden, hat über die Jahre deutlich zugenommen.<sup>9</sup> Die Anzahl der Monografien und Artikel in Sammelbänden ist recht konstant geblieben. Interessant ist, dass der Anteil grauer Literatur und technischer Reports, im Vergleich zu den Vorjahren, rückläufig ist.

<sup>9</sup> Bei den Qualifikationsarbeiten muss ebenfalls von einer erheblichen Untererfassung ausgegangen werden, da insbesondere Bachelor- und Masterarbeiten im Allgemeinen nur aufgrund der Rückmeldungen von den Datennutzern erfasst werden können und sie über Recherche schwer auffindbar sind. Hinzu kommt, dass

## Nutzungsumfang

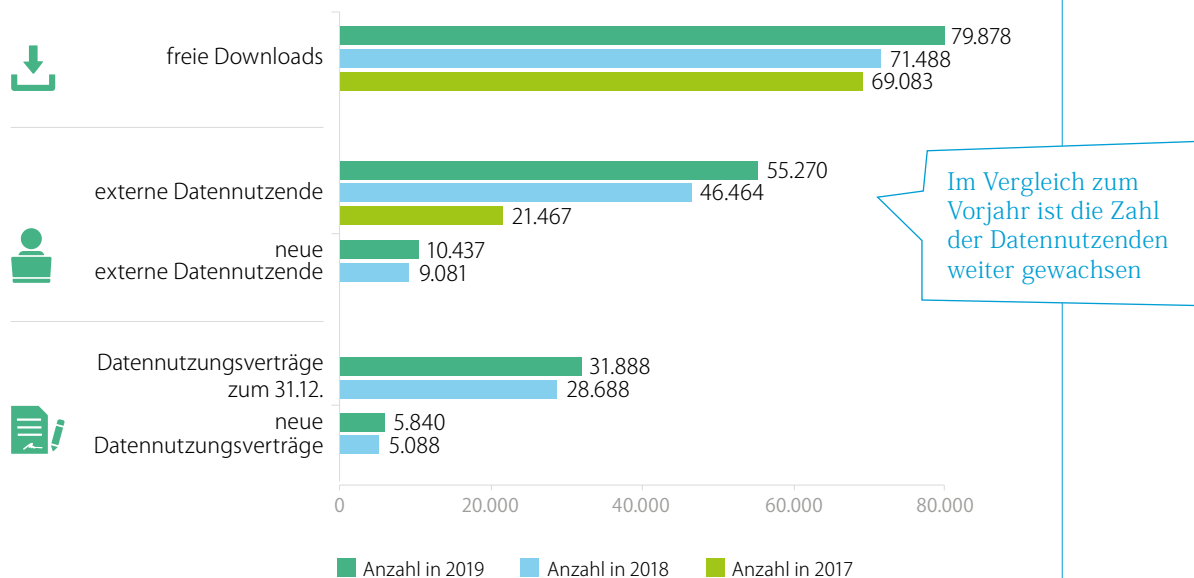
Der Nutzungsumfang stellt einen weiteren zentralen Indikator für die Bedeutung der FDZ in der Forschungslandschaft dar. Ein zentraler Wert ist dabei die Anzahl der über die FDZ bezogenen Datensätze. Daneben interessiert auch, wie viele Forschende von den Leistungen der FDZ profitieren konnten.

Da einige FDZ stark an die eigene Institution angebunden sind, andere eher eine Dienstleistungsinfrastruktur mit einer starken Außenorientierung aufweisen, ist es nicht ganz trivial, die Nutzungszahlen einzuordnen.

Zu den zentralen Aufgaben der FDZ gehört es, für die Forschung einen umfangreichen und flexiblen Datenzugang zu schaffen und diesen ständig auszubauen und zu verbessern. Für Forschende werden unter anderem aufgrund unterschiedlicher Datenschutzerfordernungen und rechtlicher Regelungen verschiedene Zugangsmöglichkeiten angeboten. Aufgrund dieser Flexibilität von Zugangsmöglichkeiten und der unterschiedlichen Erfassung von Verträgen, Projekten und Datennutzenden in den einzelnen FDZ ist es schwierig, eine präzise Zahl für den kompletten Nutzungsumfang zu ermitteln. So konnten in den Vorjahren zum Teil Doppelzählungen bei Verträgen und Projekten oder zwischen den einzelnen Zugangswegen nicht ausgeschlossen werden.

Eine solche Vielfalt empirisch abzubilden, erweist sich als komplex. Deshalb wurde seit dem Tätigkeitsbericht 2018 die Abfrage neu konzipiert und in vier Teilindikatoren differenziert: Anzahl der freien Downloads, der externen Datennutzenden, der Datennutzungsverträge und der Anzahl der davon umfassten Nutzenden bzw. Datensätze (siehe Abb. 17).

Abb. 17: Indikatoren zur Anzahl der Datennutzenden



veröffentlichte Promotionsarbeiten teilweise innerhalb größerer Projekte erstellt werden und die Datennutzung der Autorinnen und Autoren nicht in jedem Fall bei den FDZ registriert wird. Es ist also davon auszugehen, dass noch wesentlich mehr Qualifikationsarbeiten unter Rückgriff auf FDZ-Daten verfasst werden.


**79.878**

 Freie Downloads  
n=17

### Downloads

Insgesamt ist es an 17 FDZ möglich, Datensätze frei zu beziehen. Bei einem Großteil der FDZ erfordert der freie Download eine Registrierung. Einige frei verfügbare Datensätze können aber auch ohne vorhergehende Registrierung heruntergeladen werden – was die Identifikation individueller Nutzender erschwert. Die Zählung der Nutzenden kann daher nicht für alle Datenzugriffe dokumentiert werden. Teilweise kann auch aufgrund der technischen Gegebenheiten die genaue Anzahl der Downloads nicht ermittelt werden. Insgesamt wird deshalb die Anzahl der Nutzenden eher unterschätzt. Die elf FDZ, die unmittelbar zu den Nutzendenzahlen Angaben machen können, geben an, dass im Berichtsjahr 2019 frei verfügbare Datensätze 79.878-mal heruntergeladen wurden.

In der Regel ist das Forschungsdatenangebot der FDZ erst nach Vertragsabschluss oder Registrierung zugänglich.


**55.270**

 Externe Datennutzende  
n=29

**10.437**

 Neue  
externe Datennutzende  
n=27

### Externe Datennutzende

Insgesamt wurden im Berichtsjahr 55.270 externe Datennutzende der FDZ gezählt. Als externe Datennutzende werden in der Regel Datennutzende verstanden, die nicht dem jeweiligen FDZ bzw. der Institution angehören. 29 FDZ konnten Angaben zur Anzahl der externen Datennutzenden machen.

Im Jahr 2019 hat die Anzahl der externen Datennutzenden um 10.437 Personen zugenommen. Angaben zur Anzahl der im Jahr 2019 neu hinzugekommenen Datennutzenden konnten 27 FDZ machen. Das bedeutet, dass auch die Zahl der in 2019 neu hinzugekommenen Datennutzenden insgesamt höher liegen kann.


**31.888**

 Datennutzungs-  
verträge zum 31.12  
n=31

**5.840**

 Neue Datennutzungs-  
verträge zum 31.12  
n=33

### Datennutzungsverträge

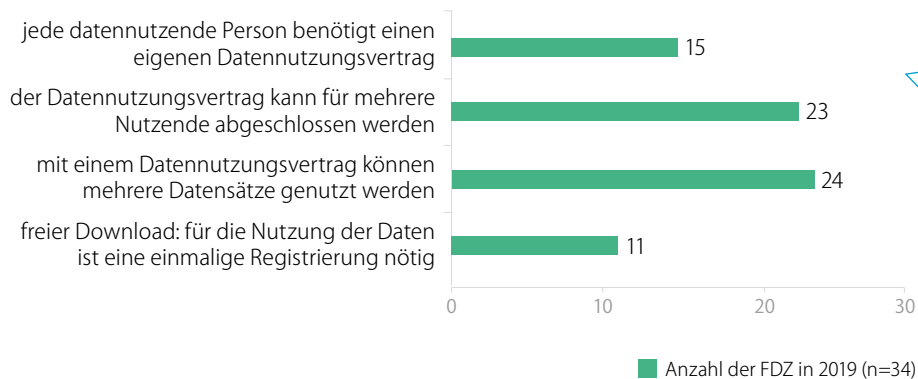
Aus Datenschutzgründen enthalten Datenbereitstellungs- und Nutzungsverträge einen expliziten Bezug zu einem Forschungsprojekt oder Qualifikationsvorhaben (Zweckbindung), d. h. für jedes Forschungsprojekt, das solche Daten nutzt, muss ein Vertrag geschlossen werden. Für die Gestaltung des Vertrages gibt es jedoch keine formalen Vorlagen. Die konkrete Vertragsgestaltung unterliegt der Vertragsfreiheit, die Regelungstiefe ist abhängig von gesetzlichen Vorgaben und Erfordernissen. So ist die gesetzliche Reglementierung bei Daten der amtlichen Statistik in der Regel recht umfassend durch Verordnungen oder Gesetze geregelt. Für viele Befragungsdaten müssen vor allem Regelungen des Datenschutzes eingehalten werden. Dies gilt insbesondere für sensible personenbezogene Informationen. Andere Daten, etwa zur Regionalisierung oder Flächennutzung, sind für bestimmte Nutzungsfälle frei verfügbar oder unterliegen bestimmten Lizenzen. Dementsprechend vielfältig zeigt sich die Vertragsgestaltung. Dies gilt nicht nur für die Daten an sich, sondern auch für die vertragsschließenden Parteien: Verträge werden mit Einzelpersonen, Projekten oder Instituten geschlossen. Teilweise umfassen sie ganze Datenbestände, Studienkollektionen oder nur einzelne Datensätze.

Am Stichtag 31.12.2019 gab es an 31 FDZ 31.888 bestehende Datennutzungsverträge in der Forschungsdateninfrastruktur des RatSWD. An 33 FDZ wurden 5.840 neue Datennutzungsverträge im Jahr 2019 abgeschlossen.

15 FDZ gaben an, dass für jede datennutzende Person ein eigener Datennutzungsvertrag abgeschlossen werden muss. 23 FDZ schließen einen Vertrag für mehrere Nutzende ab. Dabei erlauben 11 FDZ die Nutzung ausschließlich durch speziell benannte Personen. Im Durchschnitt erhielten zwei Personen über einen Vertrag Zugriffsrechte. 17 FDZ schließen projektbezogene Verträge und neun FDZ erlauben Datenzugriffe für gesamte Institute (siehe Abb. 18).

**Abb. 18: Vertragsgestaltung**

Was trifft auf Ihr FDZ zu? (Mehrfachantworten möglich)



Die Vertragsgestaltung variiert stark: Verträge können eine oder mehrere Personen oder Datensätze umfassen

24 FDZ regeln mit einem Datennutzungsvertrag den Zugriff auf mehrere Datensätze. In 13 Fällen darf nur eine benannte Person diesen nutzen, in 14 FDZ mehrere Personen. Durchschnittlich wurden dann etwas mehr als zwei Personen je Vertrag genannt. Eine projektweite Nutzung erlauben 8 FDZ. Eine Ausnahme stellt eine institutsweite Nutzung dar, die lediglich ein FDZ berichtet.

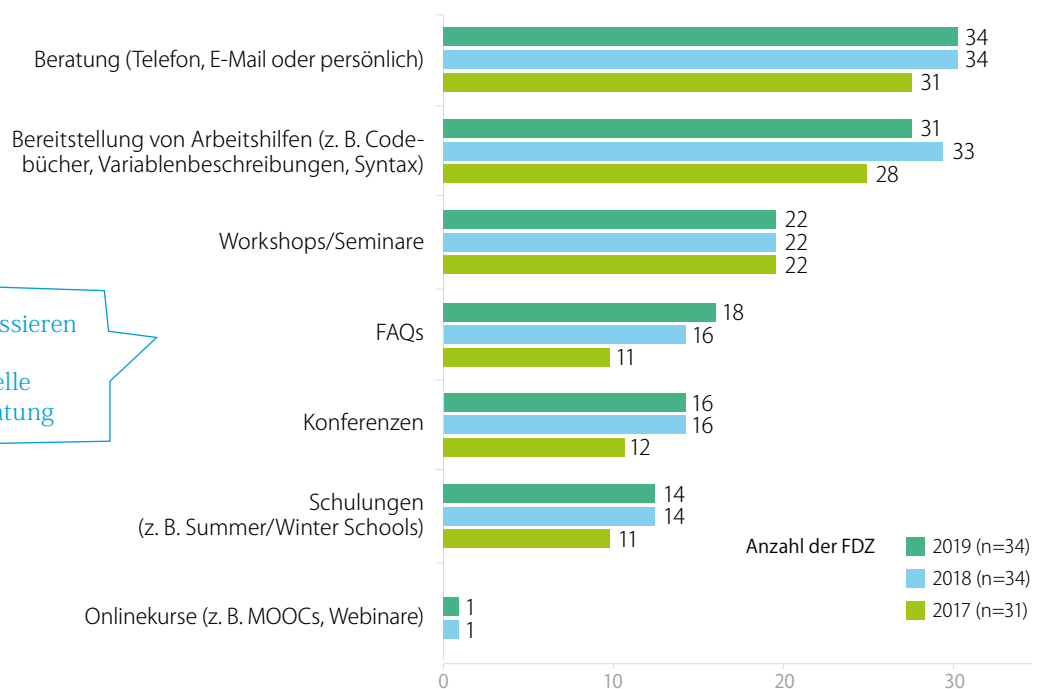
## 5 Service für Nutzerinnen und Nutzer

Dieses Kapitel stellt die nutzungsorientierte Säule der FDZ-Arbeit in den Fokus. Kapitel 2 hat bereits gezeigt, dass die Struktur der Forschungsdatenzentren einen hohen Anteil an aktiver Forschung von FDZ Wissenschaftlerinnen und Wissenschaftlern erlaubt (vgl. Abb. 5: selbstständige Forschung von wissenschaftlichem Personal, S. 16). Durch den hohen Anteil selbst forschenden Personals in den FDZ wird die Basis der Nutzungsorientierung gelegt, denn dem Personal sind so Fallstricke und Potentiale der angebotenen Daten und Servicebedarfe aus der eigenen Forschungstätigkeit heraus bekannt. Der Nutzendenservice ist ein wesentlicher Bestandteil der Arbeiten innerhalb der Forschungsdaten- und Servicezentren. Die unmittelbare Beratung der Nutzenden hinsichtlich des Datenzugangs und der Datensätze ist dabei ebenso wichtig, wie die Aufbereitung von Datensätzen und die Bereitstellung von Arbeitshilfen. Nachfolgend sollen die Nutzendenservices der FDZ näher dargestellt werden.

**Abb. 19: Nutzendenservice in den FDZ**

Welche Services werden für die Nutzenden durch Ihr FDZ angeboten?  
(Mehrfachantworten möglich)

Alle FDZ fokussieren vor allem auf eine individuelle Nutzungsberatung



© RatSWD 2020

Alle FDZ bieten weiterhin für die Belange, die von externen Forschenden an sie herangetragen werden, Ansprechpersonen, die telefonisch, per E-Mail oder direkt vor Ort kontaktiert werden können (siehe Abb. 19). Somit übernehmen die Mitarbeitenden des FDZ die verantwortungsvolle und komplexe Aufgabe, Forschende individuell zu beraten. Diese Beratung ist auch deshalb möglich, da die Mitarbeitenden in der Regel eigene Forschungserfahrungen mit den Datensätzen haben und daher mit dem Analysepotenzial vertraut sind.

Im Vergleich zum Vorjahr gibt es keine Veränderungen hinsichtlich der Anzahl der angebotenen Workshops/Seminare, Konferenzen und Schulungen. Leicht rückläufig ist die Anzahl der bereitgestellten Arbeitshilfen wie Codebücher, Variablenbeschreibungen oder Syntax; die angebotenen FAQs steigen hingegen leicht. An einem FDZ können Nutzende an einem Onlinekurs teilnehmen. Außerdem

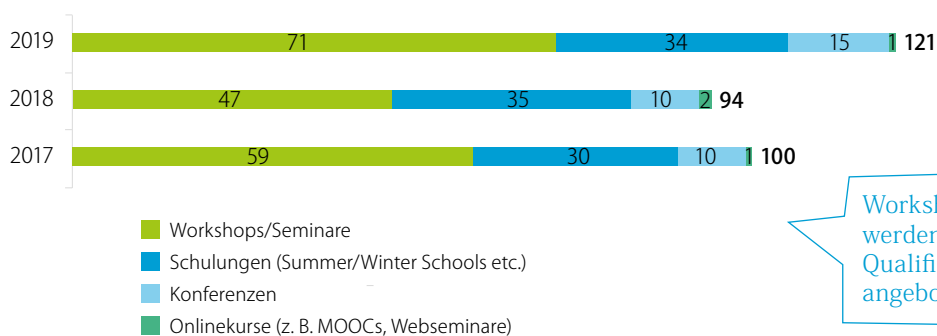
bieten vereinzelt FDZ virtuelle Hilfen in Form von Online-Tutorials/Video-Tutorials oder Diskussionsforen an. Diese Angebote sollen die Datennutzenden in die Lage versetzen, das Analysepotenzial spezifischer Datensätze besser auszuschöpfen (siehe Kapitel 6).

## Qualifizierungsmaßnahmen

Neben dem reinen Nutzendenservice bieten die FDZ eine Vielzahl von Maßnahmen zur Qualifizierung in unterschiedlichen Formaten an, die die Forschenden für die Arbeit mit den Datensätzen qualifizieren sollen. Insgesamt bieten 24 FDZ 121 Qualifizierungsmaßnahmen an und erhöhen somit nochmals ihren Beitrag zur Qualifizierung von Nutzenden um 27 Veranstaltungen. Die Durchführung von qualifizierenden Workshops ist dabei ein besonders probates Mittel, um Nutzende in relativ kleinen Gruppen zu schulen: So wurden 71 Workshops/Seminare im Berichtsjahr durchgeführt, ein Anstieg um 24 Veranstaltungen im Vergleich zum Vorjahr. Während die Schulungen und Onlinekurse konstant sind, zeichnet sich auch für die Konferenzen ein leichter Anstieg ab.

**Abb. 20: Qualifizierungsmaßnahmen**

Wie viele der folgenden Qualifizierungsmaßnahmen bietet Ihr FDZ an?  
(Mehrfachantworten möglich)

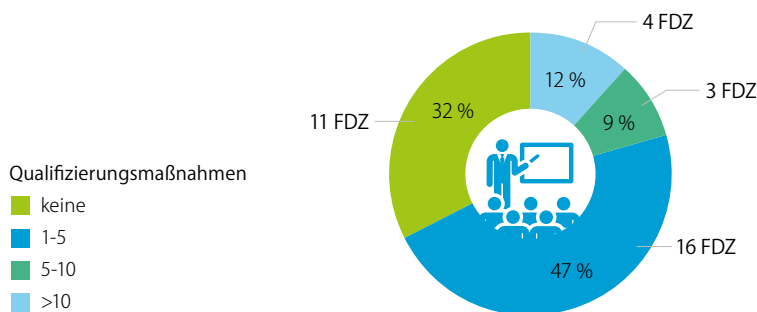


Workshops/Seminare werden am häufigsten als Qualifizierungsmaßnahmen angeboten.

Die angebotenen Qualifizierungsmaßnahmen sind dabei sehr unterschiedlich über die FDZ verteilt (Abb. 21). 11 FDZ bieten keine Qualifizierungsmaßnahmen in dieser Form an. Die Mehrzahl der akkreditierten FDZ führen zwischen einer und fünf Qualifizierungsmaßnahmen durch. Bemerkenswert ist die Vielzahl (mehr als 10 Veranstaltungen) an Qualifizierungsangeboten, die vier FDZ im Angebot haben.

**Abb. 21: Anzahl der Qualifizierungsmaßnahmen pro FDZ**

Geclustert, n=34



Die Anzahl der angebotenen Qualifizierungsmaßnahmen pro FDZ variiert stark



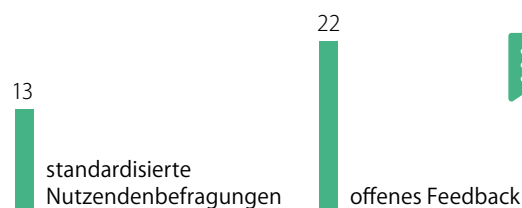
## Qualitätssicherung der Serviceangebote

Die meisten FDZ führen prozessbegleitend Maßnahmen durch, um die Qualität ihrer Serviceangebote zu sichern. Insgesamt 13 FDZ nutzen dazu standardisierte Nutzendenbefragungen, und 22 FDZ ziehen offenes Feedback der Nutzenden zur Qualitätssicherung heran. Weitere Feedbacks erheben einige FDZ über Rückmeldemöglichkeiten auf ihrer Webseite und im Rahmen von Feedback-Fragebögen im Anschluss an Schulungen. Die Häufigkeit der Durchführung der standardisierten Nutzendenbefragung unterscheidet sich deutlich. Sieben FDZ geben an, diese kontinuierlich durchzuführen, während zwei FDZ diese jährlich und vier FDZ diese seltener als jährlich durchführen. Die Ergebnisse werden von den FDZ aufgenommen, diskutiert, bewertet und wenn möglich umgesetzt. Teilweise werden die Ergebnisse auf den Seiten der FDZ veröffentlicht, dienen als Richtungsweiser für Weiterentwicklungen der Infrastruktur oder werden zur Berichterstattung genutzt. Die Ausgestaltung des offenen Feedbacks ist divers und reicht von bilateralen Gesprächen mit den Nutzenden bis hin zur Qualitätskontrolle des wissenschaftlichen Beirates des FDZ oder des angeschlossenen Instituts.

**Abb. 22: Qualitätssicherung durch Nutzendenbefragungen**

Wie wird dieses Serviceangebot evaluiert bzw. die Qualität gesichert?  
(Mehrfachantworten möglich)

■ Anzahl der FDZ in 2019 (n=22)



Qualitätssicherung findet mehrheitlich durch offenes Feedback der Nutzenden statt

© RatSWD 2020

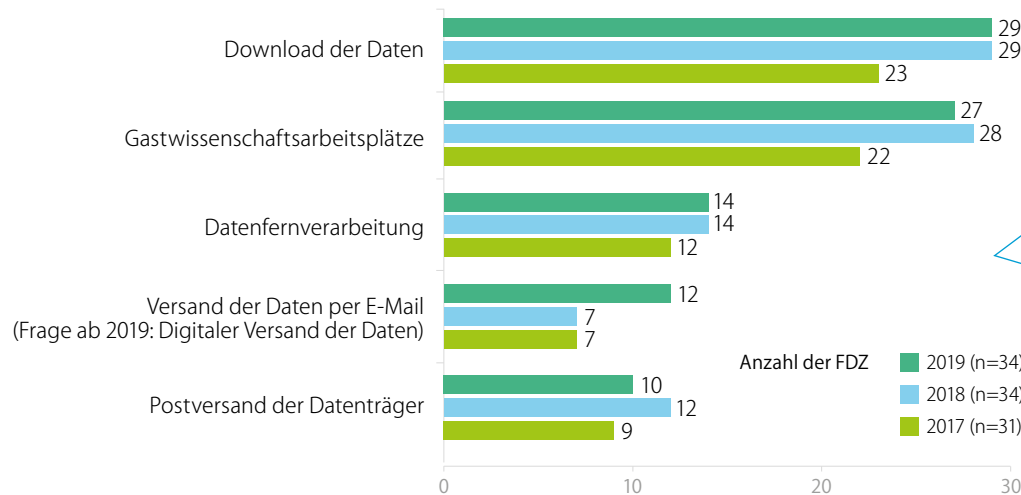
Zusammenfassend zeigt sich, dass die Häufigkeit der Qualitätssicherung ebenso unterschiedlich ist wie die Ausgestaltung. Gemeinsam ist ihnen jedoch, dass die FDZ besonderen Wert auf die Umsetzung des Feedbacks legen und versuchen, Hinweise bei der Durchführung von Veranstaltungen umzusetzen und das Serviceangebot zu optimieren, unabhängig von der Art der Qualitätssicherung.

## Zugangswege und Formate für die Datenbereitstellung

Die Bereitstellung der Daten an den FDZ erfolgt über zwei Zugangswege: (1) On-Site-Nutzung und (2) Off-Site-Nutzung. Bei der On-Site-Nutzung (1) können Forschende an Gastwissenschaftsarbetsplätzen (GWAP) auch auf vertrauliche und datenschutz sensible Daten zugreifen (in Abb. 2 ist die Verteilung der Gastwissenschaftsarbetsplätze in Deutschland dargestellt; der Infokasten 2 (S. 36) gibt Auskunft zur Definition von GWAP). (2) Bei der Off-Site-Nutzung können Datennutzende außerhalb der FDZ auf die Daten zugreifen. Je nach Datensatz kann dieser nach Download oder nach postalischem bzw. digitalem Versand – meist mit zuvor unterschriebenem Nutzungsvertrag – von den Forschenden innerhalb der eigenen Infrastruktur genutzt werden. Darunter fällt auch die Datenfernverarbeitung, etwa über VPN oder über einen Zugang zu einem Webportal.

**Abb. 23: Angebotene Datenzugangswege**

Über welche Zugangswege werden die Daten Ihres FDZ angeboten?  
(Mehrfachantworten möglich)

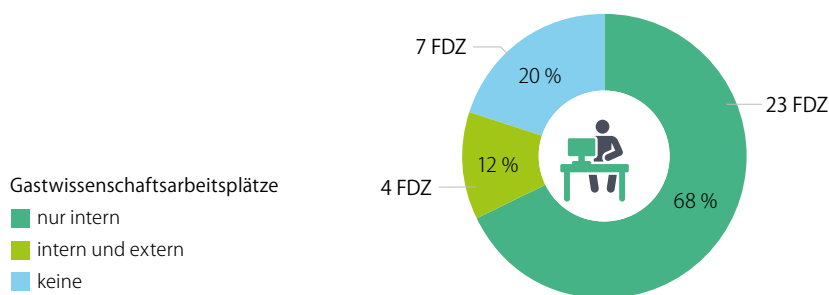


Digitaler Versand  
der Daten gewinnt  
an Bedeutung

Mehr als Dreiviertel der FDZ (27) bieten Gastwissenschaftsarbeitsplätze an. Von diesen bieten vier FDZ auch die Möglichkeit an, die Daten auf einem Gastwissenschaftsarbeitsplatz außerhalb der eigenen Institution zu verarbeiten (in Abb. 2, S. 14 ist die Verteilung von Gastwissenschaftsarbeitsplätzen in Deutschland dargestellt). Im Vergleich zum Vorjahr bietet jedoch ein FDZ keine Möglichkeit mehr, auf sensible Daten zugreifen zu können (siehe Abb. 23).

**Abb. 24: Gastwissenschaftsarbeitsplätze 2019**

Über welche Zugangswege werden die Daten Ihres FDZ angeboten? (n=34)



Gastwissenschafts-  
arbeitsplätze  
ermöglichen Zugang  
zu sensiblen Daten

Bei der Off-Site Nutzung bieten die FDZ für die Forschenden verschiedene Zugangswege an. Wie im Vorjahr haben insgesamt 29 FDZ die Off-Site-Datennutzung über einen Download der Datensätze eingerichtet. Die Datennutzung über eine (kontrollierte) Datenfernverarbeitung, z. B. über eine Remoteverbindung oder die Einsendung von Analyseskripten, bieten weiterhin 14 FDZ an. Ein Datenversand per E-Mail wird von zwölf, der Postversand eines Datenträgers von zehn FDZ angeboten. In der Regel ist dabei der digitale Versand oder der Versand per Post alternativ, das heißt FDZ versenden die Daten entweder per Post oder digital. Nur ein FDZ bedient hier beide Bereitstellungswege. Es zeigt sich, dass der Datenversand via E-Mail im Vergleich zum Vorjahr (7 FDZ) deutlich an Bedeutung gewonnen hat, während der Postversand rückläufig ist.

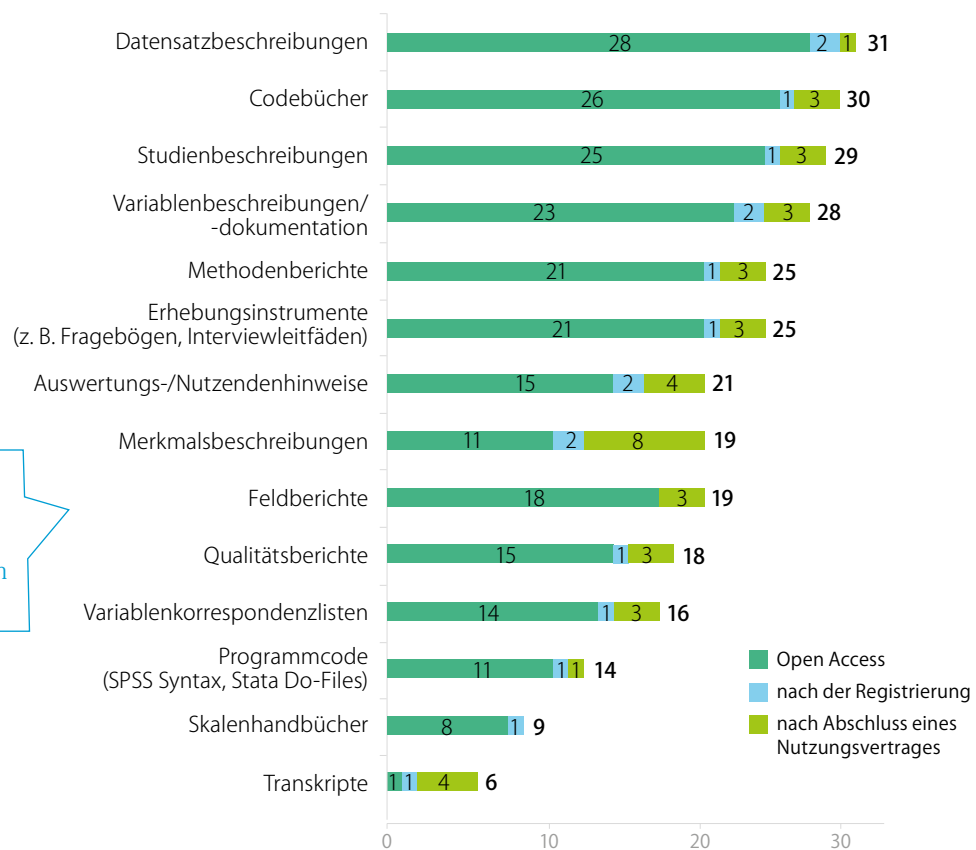
## Bereitstellung von Arbeitshilfen

Um einen guten Einblick über die angebotenen Datensätze und somit in die Umsetzung der Forschungsfragen mithilfe der Datensätze zu erlangen, ist es für Datennutzende hilfreich auf datensatzspezifische Arbeitshilfen zurückgreifen zu können. Beispielsweise können durch ausführliche Datensatzbeschreibungen oder Codebücher, Datensätze vor der Beantragung hinsichtlich ihres Mehrwertes für das Forschungsprojekt geprüft und dadurch das Beantragen von nicht benötigten Datensätzen reduziert werden. Davon profitieren somit zum einen die Nutzenden, zum anderen aber auch die FDZ selbst, da sie Arbeitsaufwand für unnötige Verträge vermeiden. Die nachfolgende Abbildung zeigt daher die verschiedenen Arbeitshilfen der FDZ, und wie diese bereitgestellt werden, auf.

**Abb. 25: Bereitstellung von Arbeitsmaterialien in 2019**

Welche Arbeitshilfen werden über welchen Zugangsweg bereitgestellt? (n=33)  
(Mehrfachantworten möglich)

Die FDZ stellen umfangreiche und vielfältige Arbeitsmaterialien zur Verfügung



© RatSWD 2020

Fast alle FDZ stellen zu jedem angebotenen Datensatz Informations- und Dokumentationsmaterialien zur Verfügung. Diese variieren je nach FDZ und Datensatz. Die Arbeitshilfen werden den Forschenden mehrheitlich Open Access zur Verfügung gestellt, sodass sich die Forschenden bereits vor dem Zugang zum Datensatz Informationen einholen können<sup>10</sup>. Am häufigsten werden Datensatzbeschreibungen, Codebücher, Studienbeschreibungen, Variablenbeschreibungen, Variablendokumentationen sowie Methodenberichte bereitgestellt.

<sup>10</sup> Open Access beschreibt den freien Zugang zu wissenschaftlicher Literatur und anderen Materialien im Internet.

Häufig werden zusätzlich datensatzspezifische Arbeitshilfen, wie etwa Programmcodes (z. B. Do-Files, Syntax-Files), technische Metadatenbeschreibungen (z. B. XML-Files), Webanwendungen (z. B. Metadatenrecherchesysteme) oder theorieorientierte Dokumente bereitgestellt, die über den inhaltlichen Hintergrund einzelner Fragen oder Variablen informieren. Insbesondere bei Programmcodes und Transkripten erfolgt der Zugang nach Abschluss des Nutzungsvertrages. Die Begründung für die Bereitstellung nach einer Registrierung oder Abschluss eines Vertrages liegt in der Art der Arbeitshilfen: Diese Arbeitshilfen enthalten z. B. sensible Informationen, die aus Datenschutzgründen nicht immer offen zugänglich sind, oder eine Syntax, die erst nach dem Vorliegen von Datensätzen nützlich ist. In manchen Fällen findet vor der Bereitstellung auch eine formale Prüfung hinsichtlich einer nicht-kommerziellen Nutzung statt.

In der Regel werden die Arbeitshilfen online über die Webseiten zum Download angeboten und können ohne eine vorherige Registrierung als Open-Access-Material bezogen werden. Werden die Forschungsdaten den Nutzenden in einem Paket zur Verfügung gestellt, so beinhaltet dieses Paket in der Regel auch die verfügbaren Arbeitshilfen.

Darüber hinaus bieten einige FDZ weitere Arbeitshilfen an. Hierzu gehören beispielsweise Zitationshinweise für die Forschungsdaten oder für im Rahmen der qualitativen Forschung erhobene Daten Zusatzinformationen zu Tafelbildern, Sitzplänen oder Interviewprotokollen. Aber auch neue Kommunikationsformen, wie Video-Tutorials zur Arbeit mit der Daten-Plattform sowie auch interaktive Metadatenportale, werden genutzt, um die Datennutzenden bei der Datenarbeit zu unterstützen.

### Zeitdauer von der Antragstellung bis zum Datenzugang

Sofern es sich um Daten handelt, die direkt über einen Online-Zugang bezogen werden können, ist der Zugriff auf die gewünschten Daten sofort bzw. binnen weniger Stunden möglich (z. B. über eine einfache Registrierung oder den Download aus dem Datenbestandskatalog, den die Nutzenden selbst vornehmen können). Für sensible, weniger stark anonymisierte Daten ist der Zugang über einen Datendownload nicht möglich.

Die Zeitdauer von der Unterzeichnung des Datennutzungsvertrags variiert in solchen Fällen zwischen den FDZ von einer Stunde bis hin zu mehreren Wochen. Längere Bearbeitungszeiten sind bei solchen Datensätzen zu erwarten, bei denen eine spezielle Aufbereitung der Daten erforderlich oder gewünscht ist oder spezielle Genehmigungen für die Datenfreigabe eingeholt werden müssen, beispielsweise für eine gesonderte Prüfung für die Nutzung der Daten im Ausland. Über die Hälfte der FDZ machen die Daten nach Vertragsunterzeichnung innerhalb einer Woche (19 FDZ) und zwölf FDZ innerhalb von zwei bis vier Wochen nachnutzbar. Lediglich ein FDZ benötigt dafür länger als vier Wochen.

**Infokasten 2:****Datenzugangswege und Datenzugangsformate****Gastwissenschaftsarbetsplätze (GWAP)**

Speziell gesicherte Arbeitsplätze in den Forschungsdatenzentren (FDZ), an denen Forschende auf mindestens formal anonymisierte Daten zugreifen können. Typische Merkmale dieser Gastwissenschaftsarbetsplätze sind zum Beispiel, dass sie keinen unkontrollierten Netzzugang besitzen und ein lokales Abspeichern der Daten nicht möglich ist.

**Datenfernverarbeitung**

Datenzugangsweg, bei dem die Forschenden die Daten im FDZ auswerten können, ohne selbst vor Ort sein zu müssen. Der Zugriff erfolgt durch die Übergabe von Analyseskripten an die FDZ. Diese Skripte schreiben die Forschenden an ihren eigenen Arbeitsplätzen (eventuell mit Hilfe eines Strukturdatensatzes, an dem sie die Programmsyntax testen können). Abhängig von der Datensensibilität und/oder Rechtsgrundlage werden die Ergebnisdateien im FDZ geprüft (Outputkontrolle), bevor sie an die Forschenden übermittelt werden. Man spricht dann von kontrollierter Datenfernverarbeitung oder Remote Access.

**Scientific Use Files (SUF)**

Forschungsdatensätze die grundsätzlich mindestens faktisch anonymisiert sind, aber dennoch ein hohes Analysepotenzial bieten.

**Campus Files (CF)**

Für die universitäre Lehre bestimmte und im Vergleich zu Scientific Use Files noch stärker anonymisierte Forschungsdatensätze.

**Public Use Files (PUF)**

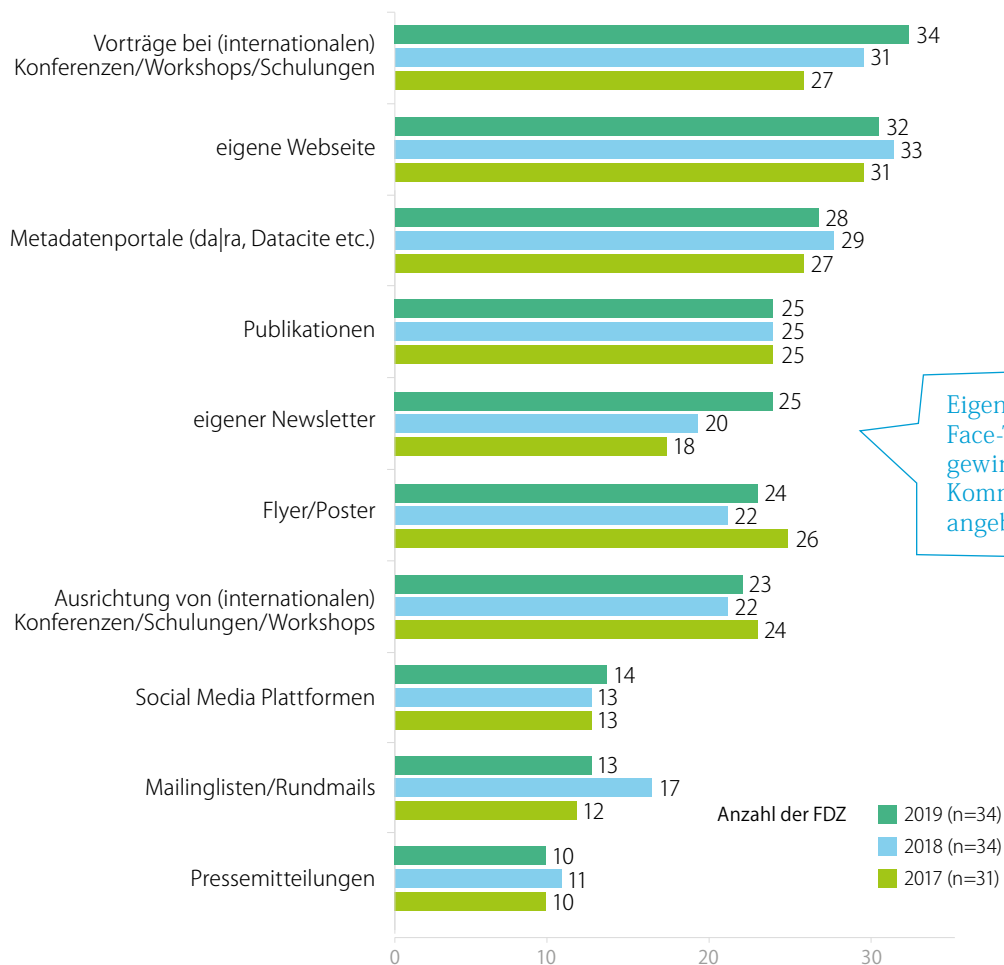
Anonyme Forschungsdatensätze, bei denen keine Nutzungseinschränkungen bestehen und somit eine Weitergabe der Daten auch außerhalb der wissenschaftlichen Forschung möglich ist.

## Bewerbung der Daten

Die FDZ sind an einer intensiven Nutzung ihrer Daten interessiert. Um den Bekanntheitsgrad des eigenen Datenangebotes zu steigern, das Forschungspotenzial aufzuzeigen und somit die Nutzung der Daten anzuregen, nutzen die FDZ verschiedene Kommunikationsformate, um ihre Datensätze zu bewerben. Diese werden im Folgenden dargestellt.

**Abb. 26: Genutzte Informationskanäle zur Kommunikation des Datenangebots durch die FDZ**

Wie und über welche Wege werden die angebotenen Daten der Wissenschaft bekannt gemacht?  
(Mehrfachantworten möglich)



Vorherrschend sind die Selbstbewerbung über die eigenen Webseiten und Newsletter der FDZ sowie Vorträge auf (internationalen) Konferenzen. Eine zentrale Rolle, um über das Datenangebot zu informieren und den Nutzendenkreis zu erweitern, spielen auch Metadatenportale wie da|ra und Datacite. Die flächendeckende Nutzung dieser Angebote stellt eine vielversprechende Basis für die Öffnung (und Auffindbarkeit) des Datenangebots auch für weitere Disziplinen auf nationaler und internationaler Ebene dar. Publikationen und das Ausrichten von Schulungen und Workshops folgen. Nach wie vor bleibt das Versenden von Pressemitteilungen sowie das Bespielen von Social Media Kanälen weniger verbreitet. Spannend wird es sein zu beobachten, wie sich diese Entwicklung im nächsten Jahr, aufgrund eingeschränkter Möglichkeiten von Präsenzveranstaltungen basierend auf der Covid-19 Pandemie, verändern wird und ob online-Bewerbungen an Bedeutung gewinnen.

## 6 Aktuelle Entwicklungen der Forschungsdateninfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften

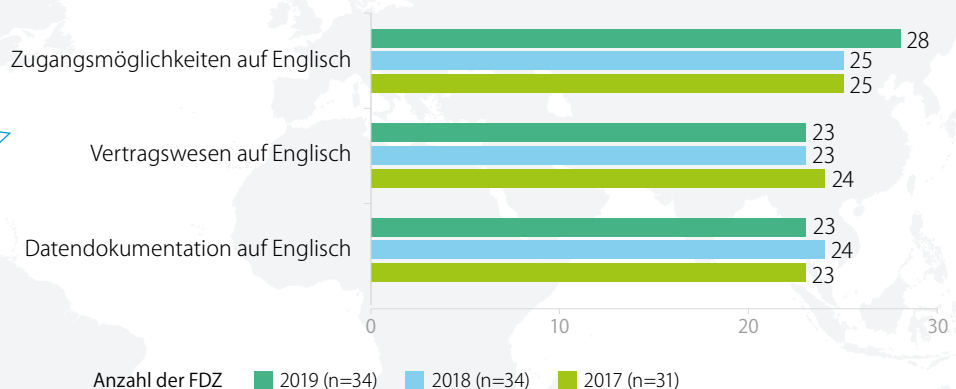
### Internationalisierung

Durch die zunehmende globale Verflechtung von Wirtschaft, Gesellschaft und Politik ergeben sich neue Forschungsfelder, die nur im Rahmen einer internationalen wissenschaftlichen Kooperation bearbeitet werden können. Die Voraussetzung dafür ist zum einen der Zugang internationaler Forscher zu den Daten in den jeweiligen Ländern. Zum anderen werden auch vermehrt internationale Mikrodaten nachgefragt, die eine bessere vergleichende Analyse erlauben.

Die weiterhin wachsende Internationalisierung der Forschungslandschaft wird auch in den meisten Forschungsdatenzentren (FDZ) bedient. Die Forschungseinrichtungen und Forschungsinfrastrukturen richten sich verstärkt international aus und reagieren damit auf die Anforderungen in der Wissenschaft. Es wurden Zugangsmöglichkeiten und Dokumentationen für die internationale Wissenschaftscommunity in Deutschland in englischer Sprache geschaffen.

**Abb. 27: Unterstützung internationaler Wissenschaftlerinnen und Wissenschaftler**  
In welcher Form unterstützen Sie Wissenschaftlerinnen und Wissenschaftler aus dem Ausland?  
(Mehrfachantworten möglich)

Drei weitere FDZ bieten Zugänge auf Englisch an

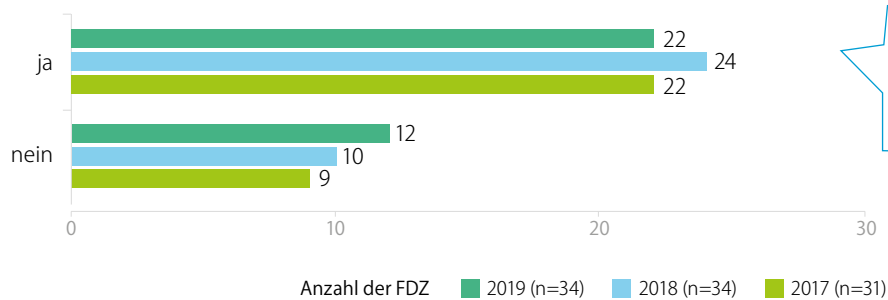


© RatSWD 2020

28 FDZ unterstützen den Datenzugang über Zugangsmöglichkeiten auf Englisch. Das sind 3 FDZ mehr als noch im Berichtsjahr 2018 bei gleichbleibender Anzahl der FDZ. Das heißt, dass die Zuwächse real sind und nicht durch neu akkreditierte FDZ zustande gekommen sind. 23 FDZ bieten Datendokumentationen auf Englisch an und 23 FDZ halten englischsprachige Verträge vor (siehe Abb. 27). Weiterhin werden die internationalen Nutzenden durch eine persönliche Beratung per E-Mail oder telefonisch unterstützt, sodass die gesamte Kommunikation auf Englisch erfolgt. Zusätzlich gibt es Arbeitsübersetzungen des Vertrages sowie Newsletter, User Workshops, Konferenzen und Schulungen in englischer Sprache. Fast alle FDZ (31) bieten dabei englischsprachige Unterstützungsmöglichkeiten, davon offerieren 27 nicht nur ein Angebot, sondern verschiedene Unterstützungen in englischer Sprache.

**Abb. 28: FDZ mit internationalen Kontakten**

Bestehen Kontakte zu ausländischen Forschungseinrichtungen?

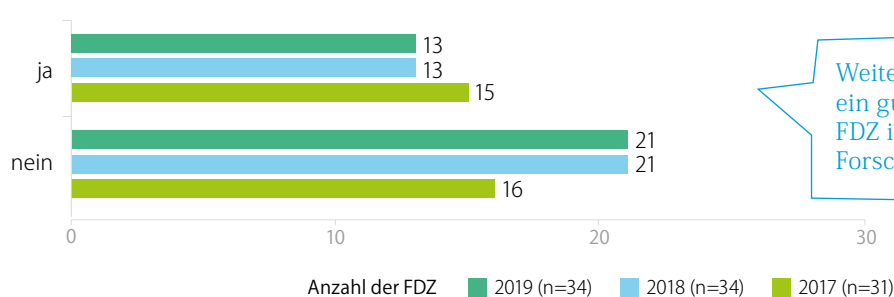


Weiterhin pflegen  
zwei von drei FDZ  
internationale  
Kontakte

Der Anteil an internationalen Datennutzenden zeigt eine große Streuung über die einzelnen FDZ hinweg. Insgesamt 22 FDZ haben Kontakte zu internationalen Forschungseinrichtungen (siehe Abb. 28, zu den Kooperationen auf nationaler Ebene siehe Abschnitt „Kooperationen und Forschungsaktivitäten“, S. 16). Hierzu zählen sowohl die Zusammenarbeit mit einzelnen Personen, internationalen Forschungsinstitutionen und -konsortien als auch die internationale Zusammenarbeit zur Abstimmung von Verfahren, so etwa auf europäischer Ebene zu Fragen des Datenzugangs oder der Geheimhaltung.

**Abb. 29: FDZ mit engen internationalen Forschungsk Kooperationen**

Gibt es internationale Forschungsk Kooperationen (d. h. wird gemeinsam an einem Thema gearbeitet)?



Weiterhin engagiert sich  
ein gutes Drittel der  
FDZ in internationalen  
Forschungsk Kooperationen

Neben der Unterstützung der internationalen Wissenschaftscommunity bei den konkreten Datennutzungen durch die Bereitstellung englischsprachiger Informationen und Materialien und vertiefend zu den allgemeinen internationalen Kontakten, arbeiten einige FDZ auch an gemeinsamen Themen im Rahmen internationaler Forschungsk Kooperationen. Für das Berichtsjahr 2019 gaben 13 FDZ an, dass sie in internationalen Forschungsk Kooperationen an gemeinsamen Themen arbeiten (siehe Abb. 29). Auch die engen angegebenen Kooperationen sind sehr vielfältig und beziehen sich überwiegend auf Projekte, Arbeitsgruppen und Zusammenarbeit mit europäischen und internationalen Einrichtungen wie Universitäten, FDZ, Datenarchiven und Forschungsinstituten.



## Innovationen und Neuerungen zur bestehenden Dateninfrastruktur

Es zeigt sich, dass die Neuerungen im Berichtsjahr 2019 sehr breit gefächert sind. Zusammenfassend nannten die meisten FDZ hier insbesondere die kontinuierliche Erweiterung des Datenangebotes, die Verfügbarmachung neuer Datensätze und -produkte sowie die Schaffung neuer Zugangsmöglichkeiten.

Auffällig ist die vermehrte Einführung neuer Datenzugangspunkte wie GWAP im In- und Ausland mit höherem Sicherheitsstandard, auch in Kombination mit Remote Access<sup>11</sup>. Somit haben die FDZ weiterhin in den Ausbau der Infrastruktur investiert. Es wurde auch vermehrt an der Verknüpfung der Datensätze und der Verbesserung der Dokumentation in Form von Metadaten, der eigenen Internetpräsenz oder der Berücksichtigung in Antragsverfahren gearbeitet.

## Weiterentwicklung der Infrastruktur

Im Folgenden werden Themen genannt, bei denen die FDZ Beratungsbedarf sehen bzw. Interesse am Wissensaustausch haben. *Datenschutz* ist ebenfalls im Berichtsjahr 2019 das Thema, bei dem die meisten FDZ Interesse am Wissensaustausch und Bedarf an Beratung zeigten. Einige weitere spezifische Punkte sind der Austausch über die Ausgestaltung von Nutzungs- bzw. Datenübergabeverträgen, verbindliche rechtliche Informationen zum Datenschutz und Urheberrecht sowie Sekundärdatennutzung und Anonymisierungsmethoden vor dem Hintergrund der Regelungen der DSGVO.

Im Bereich *Forschungsdatenmanagement* lassen sich fünf Hauptthemen identifizieren, die für die FDZ perspektivisch interessant sind:

1. Für die Bearbeitung von Nutzungsanfragen möchten sich die FDZ gerne zu den Themen Methoden und Werkzeuge der Antrags- und Vertragsverwaltung inklusive der Automatisierung von administrativen Prozessen und Archivierung austauschen. Die Entwicklung von Kriterien für die Datenqualität, Maßnahmen zur Steigerung der Sekundärnutzungen, Entwicklung von Portfolios und Tools für die Unterstützung der Datenvorbereitung im Forschungsprozess und die Qualifikation von Nachwuchskräften sowie Nutzendenschulungen und Anfragemanagementsysteme wurden ebenfalls genannt.
2. Diskussionsinteresse besteht hinsichtlich Methoden und Werkzeugen zu Datenharmonisierungen, Erleichterungen des Informationsflusses in Form von strukturierten Metadaten von der Datenerhebung bis hin zur Endnutzung.
3. Im Bereich Datenzugang möchten sich die FDZ zu den Themen alternative und innovative Datenzugangswege, Ausbau des „FDZ in FDZ“-Ansatzes, Datenfernverarbeitung, (automatisierte) Outputkontrolle und der Umsetzung von Remote Access austauschen. Gewünscht wurde ebenfalls die Definition des Begriffs der wissenschaftlichen Einrichtung mit entsprechenden Positiv- oder Negativlisten.
4. Zum Thema sensible und qualitative Daten wurde Austausch hinsichtlich der folgenden Punkte gewünscht: Datenschutzmaßnahmen bei qualitativen Daten sowie deren Erschließung und Bereitstellung sowie Remote Access für sensible Daten und die Kombination unterschiedlicher Forschungsdaten und Mixed-Methods-Ansätze.
5. Der Ausbau der Forschungsdateninfrastruktur und die Vernetzung zwischen den FDZ spielen weiterhin eine wichtige Rolle. Auch der Umgang mit neuen Aufgaben, wie zum Beispiel die Rolle von FDZ als Datentreuhänder und das Angebot von Gastwissenschaftsarbeitsplätzen für externe Daten wurde thematisiert.

*Weitere Punkte*, zu denen sich die FDZ Austausch und Beratung wünschen, waren etwa die Harmonisierung der Prozesse mit anderen FDZ, die Bereitstellung von Tools zur Datensatzsuche und zur Kommunikation mit den Nutzenden (z. B. Nutzendenportale). Zu Standards und Automatisierungsmöglichkeiten der Export-Kontrolle bei Datenfernverarbeitung und Gastwissenschaftsarbeitsplätzen besteht ebenso Beratungsbedarf wie zu Zertifizierungsverfahren und vernetzten Gastwissenschaftsarbeitsplätzen.

<sup>11</sup> In diesen GWAP ist nur die bereitgestellte IT-Ausstattung für den Zugriff auf die Mikrodaten erlaubt. Ein

## 7 Fokusthema Datenzusammenführung

Mit dem Fokusthema „Datenverknüpfung“ soll ein Überblick gegeben werden, inwieweit die Verknüpfung von Datensätzen in den vom RatSWD akkreditieren FDZ möglich und üblich ist. Mit Datenverknüpfung meinen wir das sogenannte Record Linkage, also die direkte Verknüpfung von Daten von mindestens zwei Datensätzen über eine eindeutige Identifikationsnummer (ID) beziehungsweise eine eindeutige Kombination von Merkmalen. Hierdurch können Informationen zu einzelnen Personen, Firmen oder Regionalkennziffern bzw. Geodaten, die in verschiedenen Erhebungen und/oder amtlichen Statistiken benutzt wurden, zusammengeführt werden.<sup>12</sup>

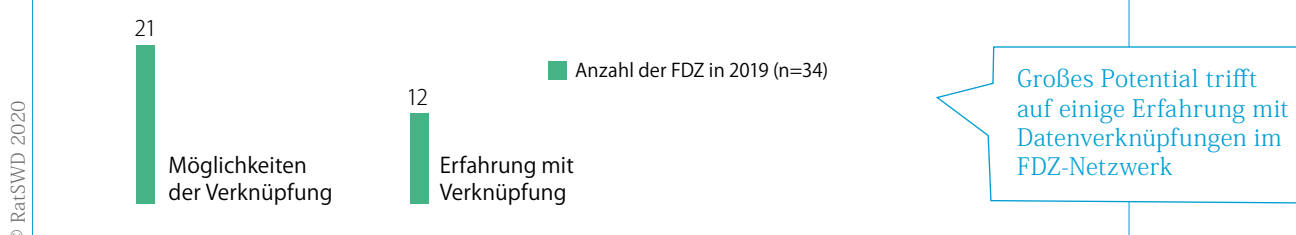
### Bedeutung und Potential des Themas Datenverknüpfung für die FDZ

Datenverknüpfung ist in vielen Forschungsdatenzentren ein wichtiges Thema, da es das Analysepotenzial einzelner Datensätze zusammenführt und damit erhöht und auf diese Weise Forschungsfragen potenziell beantwortet werden können, die aufgrund fehlender Informationen bisher nicht untersucht werden konnten.

Eine Mehrheit von 21 FDZ sieht zumindest Möglichkeiten der Verknüpfung ihres Datenbestandes. Mit 18 FDZ sind es mehr als die Hälfte der Datenzentren, die mindestens<sup>13</sup> 71 für Datenverknüpfungen geeignete Datensätze mit mindestens einem eindeutigen Identifikator oder einem Merkmalssatz im Portfolio haben. Die einzelnen Individualdaten in den Datensätzen enthalten Identifikatoren wie z. B. die Sozialversicherungsnummer, Personalausweisnummer, Handelsregisternummer, Betriebsnummer oder Steueridentifikationsnummer sowie Geodaten.

Über tatsächliche Erfahrungen mit der Verknüpfung von Daten berichten hingegen nur zwölf FDZ. Von diesen wurden in der Regel mehrere Datensätze, insgesamt über alle FDZ hinweg 29 Datensätze, verknüpft. Bei fast zwei Dritteln dieser Verknüpfungen (19 Datensätze) waren an der Verknüpfung andere FDZ beteiligt.

Abb. 30: Möglichkeiten und Erfahrungen bei der Verknüpfung von Datensätzen



Die Gründe, warum Verknüpfungen nicht noch häufiger durchgeführt werden, liegen unter anderem in den Restriktionen, die solche Projekte erschweren oder verhindern.

Zugang zum Internet und zu anderen Informationsquellen ist nicht möglich. Es gibt eine physische Zugangskontrolle und die Einhaltung der vertraglich vereinbarten Einschränkungen wird überwacht; geschultes Personal, welches sich um den Sicherheitsstandard kümmert stellt sicher, dass nur berechtigte Personen Zugang haben können (der letzte Punkt bezieht sich zum Beispiel auch auf physische Sicherheitsmaßnahmen wie Sichtschutz an Fenstern und Barrieren innerhalb des Raums, die einzelne Zugangsgesetze, d. h. Bildschirme abgrenzen). Die genaue Gestaltung des Raums ist von den Sicherheitsbedürfnissen der Datenanbieter abhängig.

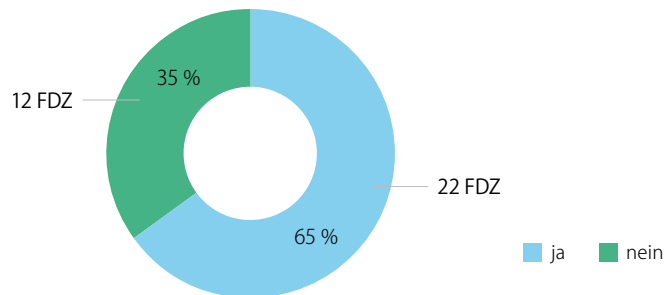
<sup>12</sup> Nicht betrachtet wird hier „Statistical Matching“, bei dem Beobachtungen aus verschiedenen Datensätzen über ein Ähnlichkeitsmaß verknüpft werden.

<sup>13</sup> Diese Zahl ist nur eingeschränkt aussagefähig, da einige FDZ jeden Datensatz auflisten, andere allgemeine Aussagen zu Typen von Datensätzen machen (z. B. „alle“ Datensätze oder allgemein die „Wirtschaftsstatistiken“ oder „Umweltstatistiken“).

**Abb. 31: Relevanz von Restriktionen**

Sind für die Daten Ihres FDZ gesetzliche oder anderen Restriktionen für die Verknüpfung von Datensätzen relevant? n=34

Fast zwei Drittel der FDZ sehen Schwierigkeiten bei der Verknüpfung von Daten



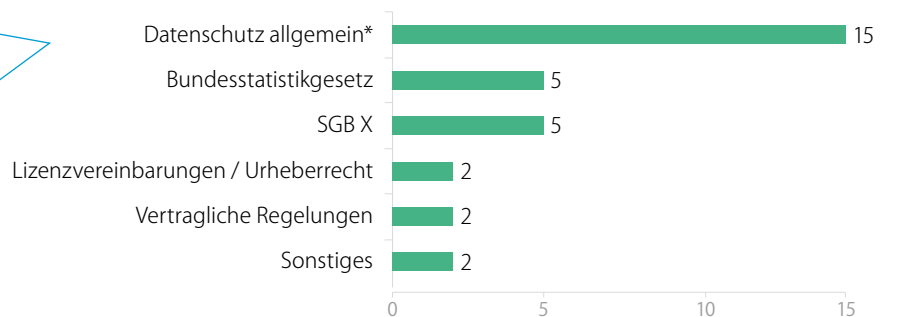
© RatSWD 2020

22 FDZ und damit fast zwei Drittel geben an, dass die Datenverknüpfungen Restriktionen unterworfen sind, die zumeist gesetzlicher Natur sind.<sup>14</sup> 15 FDZ verweisen dabei auf Datenschutzregelungen der europäischen Datenschutzgrundverordnung (DSGVO) sowie der Datenschutzgesetze des Bundes (BDSG) oder der Länder. Diese sind immer dann zu beachten, wenn es sich um personenbezogene oder -beziehbare Daten handelt, denn deren Verarbeitung und Zusammenführung bedarf stets einer rechtlichen Grundlage, wie z. B. einer informierten Einwilligung der Studienteilnehmenden oder anderer Erlaubnistatbestände. Jeweils fünf Mal werden außerdem im Speziellen das Bundesstatistikgesetz und das Sozialgesetzbuch (SGB X) als gesetzliche Beschränkungen genannt. Nur für jeweils zwei FDZ ergeben sich die Beschränkungen aus vertraglichen Vereinbarungen und aus Lizenzvereinbarungen bzw. dem Urheberrecht (siehe Abb. 32).

**Abb. 32: Relevante Restriktionen, die von außen gegeben werden**

Welche gesetzlichen oder anderen Restriktionen sind für die Daten Ihres FDZ bei der Verknüpfung von Datensätzen relevant? (Auswertung offene Frage, Mehrfachnennungen möglich)

Record Linkage ist in den FDZ derzeit hauptsächlich datenschutzrechtlich reguliert



\* entweder allgemeine Antwort „Datenschutz“ oder DSGVO, Bundesdatenschutzgesetze, Landesdatenschutzgesetze

■ Anzahl der FDZ in 2019 (n=22)

© RatSWD 2020

Das bedeutet, jemand könnte eventuell eine Verknüpfung vornehmen, aber das befragte Forschungsdatenzentrum darf dies nicht. Damit sind Datenverknüpfungen offenbar starke Grenzen gesetzt, die zwar einzelne Projekte gelegentlich möglich machen, aber vermutlich vielen denkbaren Projekten entgegenstehen.

<sup>14</sup> Die folgenden Angaben beruhen auf Auswertung einer offenen Frage, d. h. Antworten wurden im Nachhinein kategorisiert.

## Ebenen der Datenverknüpfung

Verknüpft werden können Daten auf der Personen-, Betriebs- oder der Regionalebene. Nachfolgend werden zunächst die Verknüpfungen auf Personen- und Betriebs- bzw. Unternehmensebene behandelt, weil sie datenschutzrechtlich ähnliche Herausforderungen nach sich ziehen. Danach wird auf die Verknüpfung mit Regionaldaten eingegangen, die andere Verfahren erfordern, aber datenschutzrechtlich im Wesentlichen unbedenklich sind.<sup>15</sup>

### Verknüpfung von Individualdaten wie Personen- und Firmendaten

Nach der Erhebung von Forschungsdaten der Sozial-, Verhaltens- und Wirtschaftswissenschaften enthalten diese bei Weitergabe für Sekundärauswertungen im Regelfall keine Namen, Adressen und sonstige eindeutige Identifikatoren, die erlauben, die Forschungssubjekte zu reidentifizieren. Dies wird den Befragten entweder bei Umfragen und im Rahmen von Interviews zugesichert oder bei prozessproduzierten amtlichen Daten im Rahmen der Anonymisierung ohnehin sichergestellt. Ein weiteres Merkmal empirischer Datensätze kann sein, dass diese pseudonymisiert werden. Auch bei diesen Daten ist eine Verknüpfung möglich.

Wenn Datensätze verknüpft werden sollen, müssen hierfür allerdings dieselben Identifikatoren verwendet werden, damit die Daten zusammengeführt werden können. Teil der Projektplanung ist es daher, sich auf mindestens einen gemeinsamen Identifikator zu einigen. Nicht alle Identifikatoren stehen allen FDZ zur Verfügung, die Verwendung der Versicherungsnummer der Sozialversicherung ist zum Beispiel nicht gestattet, wenn eine der beteiligten Institutionen nicht auf dem Gebiet der sozialen Sicherung tätig ist und die Verwendung nicht gesetzlich ausnahmsweise gestattet ist (§ 18 f. SGB IV).

Als Identifikatoren, auf deren Grundlage Daten auf diesen Ebenen zusammengeführt werden können, sind diese in den den FDZ vorliegenden Datensätzen beispielsweise in Form von Angaben zu Individuen, Unternehmen, Betrieben und zu Haushalten vorhanden.

10 FDZ verfügen über mindestens 30 Datensätze<sup>16</sup>, die auf dieser Ebene verknüpfbar wären. Davon könnten Verknüpfungen in 4 FDZ auf der Betriebsebene erfolgen, wofür die Betriebsnummer zur Verfügung stünde. Für 7 FDZ käme eine Verknüpfung über Personen in Frage, beispielsweise über die – z. T. pseudonymisierte – Sozialversicherungsnummer, die in 5 FDZ für mindestens 19 Datensätze vorliegt.

Identifikatoren auf Haushaltsebene (bzw. Bedarfsgemeinschaft) bieten nur sehr selten (zwei Nennungen) eine Möglichkeit der Verknüpfung.

Real durchgeführt wurden Verknüpfungen von Daten über Merkmale, die zu einer Person gehören, von sechs FDZ, und zwar entweder über die Sozialversicherungsnummer (3 FDZ für 15 Datensätze), über eine Merge-ID (zwei FDZ für drei Datensätze) oder in einem weiteren FDZ über ein Merkmalsset (Namens-/Adressdaten, Verknüpfung für zwei Datensätze). Auf der Betriebs-/Unternehmensebene wurden von zwei FDZ insgesamt vier Datensätze verknüpft, davon zwei über die Betriebs- bzw. Unternehmensnummer.

Damit ist der wichtigste Identifikator das Individuum, an zweiter Stelle folgt dann der Betrieb oder das Unternehmen.

<sup>15</sup> Ein Datenschutzproblem kann nur dann auftreten, wenn für einzelne Regionen geringe Fallzahlen bestimmter Personengruppen vorliegen, die eine Identifizierung einzelner Personen unter Umständen möglich machen.

<sup>16</sup> Die Anzahl der Datensätze ist nur eingeschränkt aussagefähig, da einige FDZ jeden Datensatz auflisten, andere allgemeine Aussagen zu Typen von Datensätzen machen (z. B. „alle“ Datensätze oder allgemein die „Wirtschaftsstatistiken“ oder „Umweltstatistiken“).

### Verknüpfung von Daten über Geodaten und andere regionale Identifikatoren

Von 9 FDZ werden 31 Projekte genannt, bei denen Daten auf Regionalebene verknüpft werden könnten. Zwei Drittel dieser FDZ (6 FDZ) geben an, dass Verknüpfungen ihrer Datensätze über raumbezogene Angaben möglich seien. Georeferenzen werden hierbei am häufigsten genannt und liegen als eindeutige IDs auch für mindestens 29<sup>17</sup> Datensätze vor.<sup>18</sup> Dabei wird erkennbar, dass einige FDZ sehr viele Verknüpfungsprojekte über Regionalinformationen durchführen könnten. In vielen Fällen wird die Georeferenz als mögliche Ebene der Verknüpfung genannt, in anderen Nennungen weitere regionale Identifikatoren (wie z. B. Adressen, Gemeinde, Kreise, Raumordnungsregionen, Staat).

Tatsächlich wurden Verknüpfungen über raumbezogene Identifikatoren bislang nur sehr selten durchgeführt. Nur ein FDZ hat bislang drei Datensätze über Georeferenzen zusammengeführt.

### Berufe und Schulen

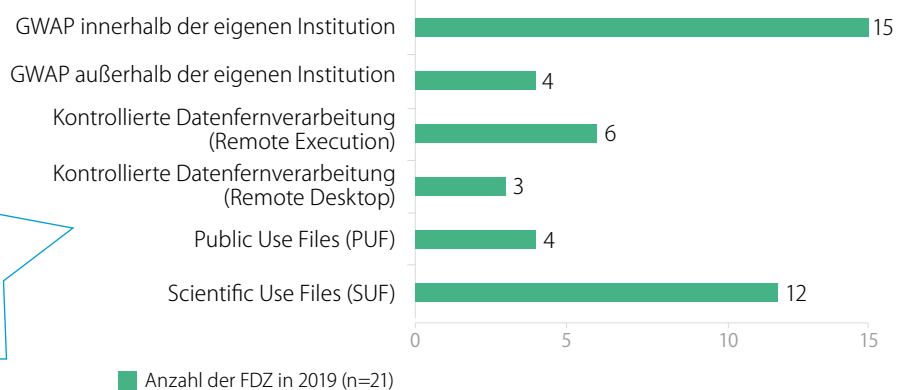
Es gibt auch Nennungen, die auf die Möglichkeiten einer Verknüpfung auf Berufsebene und auf der Ebene von Schulen hinweisen. Diese Verknüpfungen spielen in der Menge der möglichen für Verknüpfungen geeigneten Datensätze jedoch eine untergeordnete Rolle und wurden bislang auch noch nicht durchgeführt.

### Zugangswege für Verknüpfungen

Für die Verknüpfung von Datensätzen stehen in den FDZ unterschiedliche Zugangswege zur Verfügung.

**Abb. 33: Vorhandene Zugangswege**

Welche Zugangswege bietet Ihr FDZ aktuell für Verknüpfungen?  
(Mehrfachantworten möglich)



Am häufigsten werden von FDZ zur Verknüpfung von Daten Gaswissenschaftsarbeitsplätze innerhalb der eigenen Institution zur Verfügung gestellt

<sup>17</sup> Die Anzahl der Datensätze ist nur eingeschränkt aussagefähig, da einige FDZ jeden Datensatz aufführen, andere allgemeine Aussagen zu Typen von Datensätzen machen (z. B. „alle“ Datensätze oder allgemein die „Wirtschaftsstatistiken“ oder „Umweltstatistiken“).

<sup>18</sup> Als weitere IDs werden viermal Kreis/Gemeindekennziffern sowie zweimal Regionalkennzeichen genannt. Weitere jeweils einmal genannte IDs sind z. B.: Inspire Grid, Adresse, Regionalkennzeichen, Raumordnungsregionen.

Von den 21 Forschungsdatenzentren, die angeben, Datensätze mit mindestens einem eindeutigen Identifikator zu haben, bieten 15 FDZ diese verknüpfbaren Datensätze über Gastwissenschaftsarbeitsplätze in der eigenen Institution an (siehe hierzu auch die Erläuterung zu Zugangswegen und Datenformaten, Infokasten S. 36), auch vier FDZ mit Gastwissenschaftsarbeitsplätzen außerhalb der eigenen Institution öffnen diesen Weg für Datenverknüpfungen (siehe auch Abschnitt „Zugangswege und Formate für die Datenbereitstellung“, S. 32). Die Bereitstellung über GWAP ist damit auch der häufigste Zugangsweg. Immerhin 12 FDZ bieten die Datensätze als Scientific-Use-Files an, die von den Forschenden an ihrem eigenen Arbeitsplatz analysiert werden können. Nur 4 von 21 FDZ bieten die Daten auch als Public-Use-File, zum Beispiel für die Lehre an Universitäten, an.

Über den Weg des kontrollierten Fernrechnens werden Daten für Verknüpfungen von sechs FDZ im Remote Executive- und von drei im Remote-Desktopverfahren bereitgestellt; insgesamt sieben FDZ bieten Datenfernverarbeitung an. Die Bereitstellung der Daten über Gastwissenschaftsarbeitsplätze und die kontrollierte Datenfernverarbeitung tragen den besonderen Ansprüchen an die Datennutzung Rechnung: Denn die Daten verfügen über eindeutige Identifikatoren, die deshalb vielfach auf bestimmte Merkmalsträger beziehbar sein können und entsprechend sensibel sind.

Im Ergebnis bieten die verknüpften Daten in vielen Fällen einen reichhaltigen Informationsgehalt, aber sie sind oft nur an Gastwissenschaftsarbeitsplätzen zugänglich. Momentan spielt die Praxis der Zusammenführung von Daten an den FDZ jedoch (noch) eine untergeordnete Rolle. Die besonderen Anforderungen an eine Verknüpfung von Daten durch die Einhaltung von Vertraulichkeit für die Merkmalsträger und die Einhaltung legaler Rahmenbedingungen stehen dabei einer Ausweitung von Datenverknüpfungen auch zukünftig eher entgegen. Eine Tendenzaussage, ob der Zugang im Lauf der Zeit vielfältiger wird, wenn sich beispielsweise die Nachfrage positiv entwickelt, lässt sich aufgrund der erstmaligen Befragung nicht treffen.

## 8 Beschwerdemanagement

Die Sicherung und Verbesserung der Arbeitsqualität der FDZ gehört zu den Kernaufgaben des RatSWD. Bereits in der Vergangenheit war er Ansprechpartner für Beschwerden, die sich auf das Datenangebot und den Service der FDZ bezogen.

Neben der Durchführung des jährlichen Monitoringverfahrens, dessen Nebenergebnis unter anderem der vorliegende Tätigkeitsbericht ist, hat die Monitoringkommission auch die Aufgabe, ein Anlaufpunkt für Beschwerden von Datennutzenden zu sein.

Um das Beschwerdemanagement zu professionalisieren und transparent zu gestalten, wurde eine Beschwerdestelle bei der Geschäftsstelle des RatSWD eingerichtet. Sie stellt im engen Austausch mit der Monitoringkommission eine schnelle und professionelle Reaktion auf Beschwerden sicher und zieht daraus Anregungen für die Weiterentwicklung der Dateninfrastruktur.

Stellen Datennutzende gravierende Mängel im Angebot eines akkreditierten FDZ fest, so sollten sie diese beim betreffenden FDZ direkt ansprechen und versuchen, eine Lösung zu finden. Wird keine Einigung erzielt, kann das Anliegen an die Beschwerdestelle des RatSWD gerichtet werden.

Die Zuständigkeit der Beschwerdestelle beschränkt sich auf die Prüfung der Einhaltung der Akkreditierungskriterien. Der RatSWD ist nicht Ansprechpartner bei Verzögerungen im normalen Ablauf oder in Bezug auf Verhaltensweisen des Personals der FDZ. Für Beschwerden dieser Art sollte das jeweilige FDZ direkt kontaktiert werden.

Detaillierte Informationen zum Verfahrensweg finden sich im RatSWD Output 8 (5)<sup>19</sup> bzw. in aktualisierter Form auf der Webseite des RatSWD<sup>20</sup>.

### Aktuelle Beschwerdeverfahren des Berichtsjahres 2019

Im Berichtszeitraum 2019 wurde keine Beschwerde an den RatSWD gerichtet. Ein Beschwerdeverfahren aus dem Jahr 2018 zum Datenzugang bei den Statistischen Landesämtern wurde jedoch weiterverfolgt: Die Statistischen Ämter des Bundes und der Länder können nach einer Gesetzesänderung seit 2017 formal anonymisierte Einzelangaben an Gastwissenschaftsarbeitsplätzen für die Wissenschaft zugänglich machen. Für bayerische Daten wurde von dieser neuen gesetzlichen Grundlage nur partiell Gebrauch gemacht; wichtige und stark nachgefragte Daten sind weiterhin nur über die – die Forschung zeitlich und organisatorisch stark einschränkende – Nutzung der kontrollierten Datenfernverarbeitung verfügbar.

Das Beschwerdeverfahren wurde insofern als unbegründet mit der Feststellung abgeschlossen, dass es durchaus einen Zugangsweg zu den bayerischen Daten gibt (Kontrollierte Datenfernverarbeitung). Der RatSWD sieht die daraus entstehende Einschränkung dennoch als relevant für den Wissenschaftsstandort Deutschland an. Er hat daher den persönlichen und schriftlichen Austausch mit dem Bayerischen Landesamt für Statistik und mit dem Bayerischen Staatsministerium des Innern gesucht. Der Prozess auf Ebene des RatSWD hält an.

<sup>19</sup> RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017): Qualitätssicherung der vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ). RatSWD Output 8 (5). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). doi: [10.17620/02671.4](https://doi.org/10.17620/02671.4)

<sup>20</sup> <https://www.ratswd.de/beschwerdestelle>

## Anhang

---

Anhang A: Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD . . . . .	48
Anhang B: Index und Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD. . . . .	52
Anhang C: Die Monitoringkommission . . . . .	58
Anhang D: Mitwirkende bei der Erstellung des Tätigkeitsberichtes 2019 . . . . .	59



## Anhang A

### Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD

Stand: Mai 2020

Die FDZ Statistisches Bundesamt, Statistische Ämter der Länder, GML, IZA, BA im IAB und RV sind bereits vor der Gründung des RatSWD eingerichtet worden und gehörten ab 2004 zur Forschungsdateninfrastruktur des RatSWD. Bei diesen FDZ wurde das Jahr der Gründung angegeben. Alle anderen FDZ wurden nach 2004 beim RatSWD akkreditiert. Bei diesen FDZ ist jeweils das Jahr der Akkreditierung angegeben.

Gründung folgender FDZ:



1999

2001

2002

2003

2004

Gründungsausschuss

1.

Empfehlung der „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI) zur Einrichtung von Forschungsdatenzentren (FDZ)

Konstituierung des Gründungsausschusses des RatSWD)

Gründung des RatSWD

Akkreditierung folgender FDZ:



□ Anzahl der akkreditierten FDZ

2006

8  
2008

13  
2009

18  
2010

Berufsperiode

2. Berufsperiode

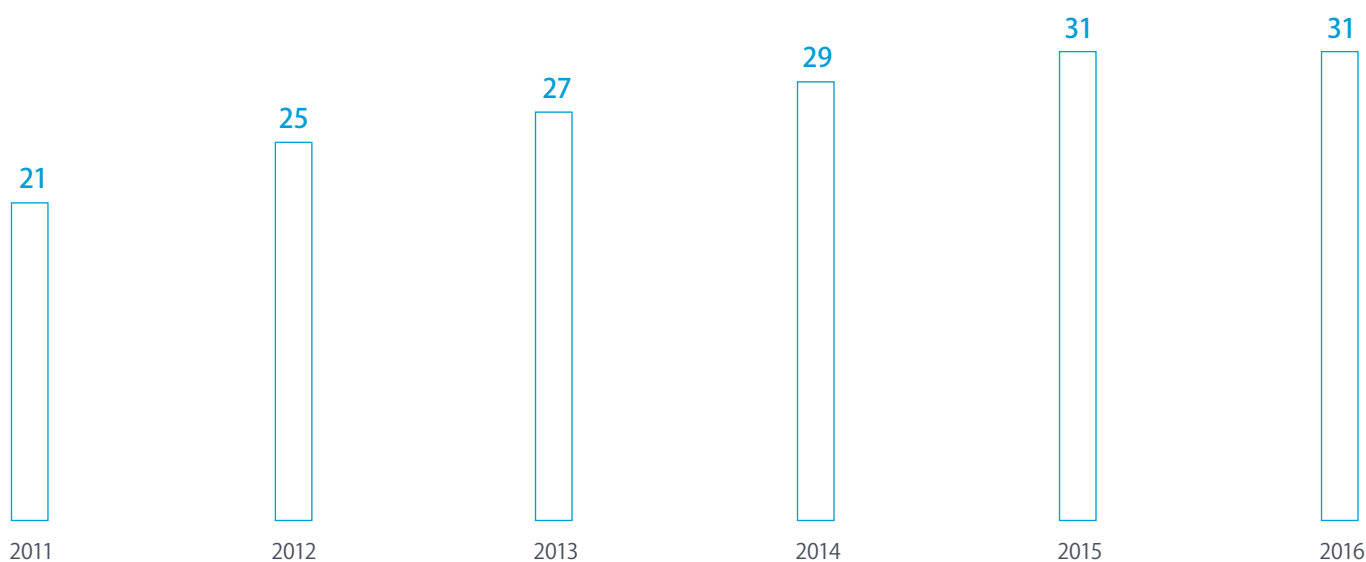
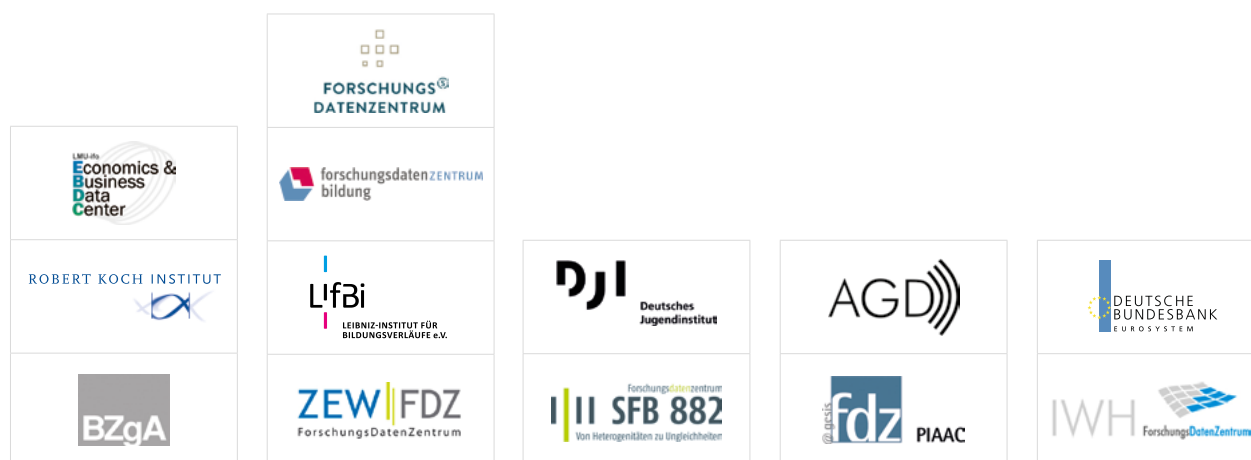
3. Berufsperiode

8 akkreditierte FDZ  
im FDI Ausschuss

Einrichtung des  
Ständigen Ausschusses  
Forschungsdaten-  
infrastruktur (FDI  
Ausschuss) des RatSWD

Einführung von  
Akkreditierungs-  
kriterien und  
Mindeststandards

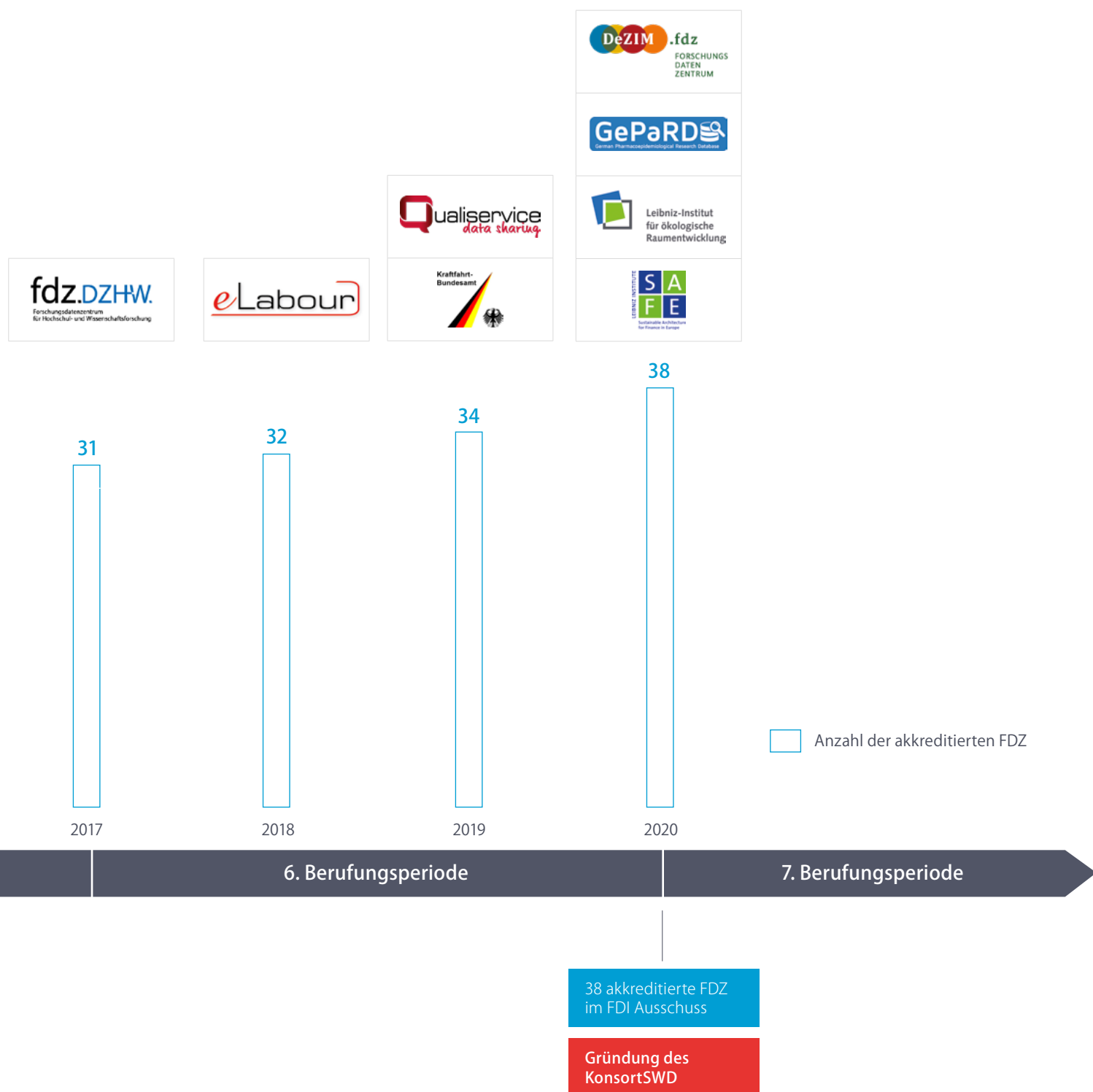
Akkreditierung folgender FDZ:



4. Berufungsperiode

5. Berufungsperiode

Einführung eines  
jährlichen Monitorings



## Anhang B

### Index und Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD

Stand: 21.07.2020

Datenangebot: ■ Sozial ■ Wirtschaft ■ Bildung ■ Gesundheit ■ Psychologie ■ Qualitativ ■ Weitere

#### BIBB-FDZ



#### Forschungsdatenzentrum im Bundesinstitut für Berufsbildung

Personen- und Betriebsdatensätze der Berufsbildungsforschung, die sich inhaltlich hauptsächlich mit dem Erwerb und der Verwertung von beruflichen Kenntnissen und Kompetenzen beschäftigen.

<https://www.bibb.de/de/53.php>



#### DeZIM.fdz

(vorläufig  
akkreditiert)



#### Forschungsdatenzentrum Deutsches Zentrum für Integration und Migration

Das DeZIM.fdz organisiert den Zugang zu Forschungsdaten, die am Deutschen Zentrum für Integrations- und Migrationsforschung erhoben werden. Zusätzlich bietet das DeZIM.fdz einen umfassenden Support zu diesen Daten und zu verschiedenen methodischen Schwerpunktthemen.

<https://dezim-institut.de/forschungsdatenzentrum-dezimfdz>



#### EBDC



#### LMU-ifo Economics & Business Data Center

Datensätze deutscher Unternehmen: Dazu zählen die vom ifo Institut erhobenen Wirtschaftsdaten zur Konjunkturentwicklung, zur Innovationsfreudigkeit und zum Investitionsverhalten sowie externe Daten zur Unternehmensfinanzierung und zur Governance-Struktur von Unternehmen.

<https://www.ifo.de/ebdc>



#### FDSZ

#### Bundesbank



#### Forschungsdaten- und Servicezentrum der Bundesbank

Verschiedene Datensätze zu Banken, Wertpapieren, Investmentfonds und Unternehmen sowie Kombinationen daraus; Studie zum Vermögen privater Haushalte.

<https://www.bundesbank.de/de/bundesbank/forschung/fdsz>



#### FDZ AGD



#### Forschungsdatenzentrum Archiv für Gesprochenes Deutsch am Institut für Deutsche Sprache

Daten zu gesprochenem Deutsch in Interaktionen (Gesprächskorpora) sowie inländische und ausländische Varietäten des Deutschen (Variationskorpora).

<http://agd.ids-mannheim.de>



#### FDZ ALLBUS



#### Forschungsdatenzentrum ALLBUS bei GESIS

Daten aus der allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) zu Einstellungen, Verhaltensweisen und Sozialstruktur der Bevölkerung in Deutschland.

<https://www.gesis.org/allbus>



#### FDZ BA im IAB



#### Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung

Daten zu Personen, Haushalten und Betrieben sowie kombinierte Datensätze aus Befragungsdaten und administrativen Daten im Bereich der Sozialversicherung und der Arbeitsmarkt- und Berufsforschung.

<https://fdz.iab.de>



**FDZ Bildung**
**Forschungsdatenzentrum Bildung am DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation**


Die angebotenen Daten aus der qualitativen Bildungsforschung umfassen AV-Daten, Transkripte und Kontextmaterialien. Im Bereich Erhebungsinstrumente der quantitativen Bildungsforschung werden Test, Fragebögen und Skalenkollektionen bereitgestellt. Die Daten und Instrumente entstammen größtenteils der Forschung zu Schul- und Unterrichtsqualität.

<https://www.fdz-bildung.de>

**FDZ-BO**
**Forschungsdatenzentrum Betriebs- und Organisationsdaten**


Quantitative und qualitative Betriebs- und Organisationsdaten, verknüpfte Personen- und Betriebsdaten sowie Daten aus Mitarbeitenden- und Mitgliederbefragungen.

<http://www.fdz-bo.diw.de>

**FDZ-Bund**
**Forschungsdatenzentrum des Statistischen Bundesamtes**


Flächendeckender Zugang zu amtlichen Mikrodaten aus den Bereichen: Bevölkerung, Bildung, Gesundheit, Wirtschaft, Agrar, Umwelt, Energie, Rechtspflege, Finanzen und Steuern.

<https://www.forschungsdatenzentrum.de>

**FDZ BZgA**
**Forschungsdatenzentrum der Bundeszentrale für gesundheitliche Aufklärung**


Daten aus regelmäßig wiederholten bundesweiten Repräsentativbefragungen zur Erreichbarkeit der Bevölkerung mit Aufklärungs- und Präventionsmaßnahmen sowie zum Wissen, den Einstellungen und dem Gesundheitsverhalten der Bevölkerung bezüglich der Gesundheitsthemen der BZgA.

<https://www.bzga.de/forschung/studien-untersuchungen>

**FDZ-Länder**
**Forschungsdatenzentrum der Statistischen Ämter der Länder**


Flächendeckender Zugang zu amtlichen Mikrodaten aus den Bereichen: Bevölkerung, Bildung, Gesundheit, Wirtschaft, Agrar, Umwelt, Energie, Rechtspflege, Finanzen und Steuern.

<https://www.forschungsdatenzentrum.de>

**FDZ-DJI**
**Forschungsdatenzentrum des Deutschen Jugendinstituts**


Daten aus den seit 1988 regelmäßig durchgeführten Surveys zum Aufwachen von Kindern und Jugendlichen sowie zu Lebenslagen von Erwachsenen und Familien.

<https://www.dji.de/abt2>

**FDZ-DZA**
**Forschungsdatenzentrum des Deutschen Zentrums für Altersfragen**


Daten der Langzeitstudien „Deutscher Alterssurvey (DEAS)“, zum Wandel der Lebenssituationen und Altersverläufe der Bevölkerung in der zweiten Lebenshälfte sowie des Deutschen Freiwilligensurveys (FWS) - eine repräsentative Bevölkerungsumfrage zum freiwilligen Engagement in Deutschland.

<https://www.dza.de/fdz.html>

**fdz.DZHW**
**Forschungsdatenzentrum für Hochschul- und Wissenschaftsforschung**


Quantitative und qualitative Daten der Hochschul- und Wissenschaftsforschung, insbesondere Studienberechtigtenpanel, Absolventenpanel, Sozialerhebung und Wissenschaftsbefragung des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW).

<https://fdz.dzhw.eu>

Datenangebot: ■ Sozial ■ Wirtschaft ■ Bildung ■ Gesundheit ■ Psychologie ■ Qualitativ ■ Weitere

#### FDZ eLabour



#### Forschungsdatenzentrum eLabour

Qualitative Daten aus Studien zur Arbeits- und Industriosozologie mit dem Schwerpunkt Wandel von Arbeit, u. a. offene bis halbstandardisierte Interviews, Beobachtungen, sowie Expertengespräche.

<http://elabour.de>



#### FDZ GePaRD

(vorläufig  
akkreditiert)



#### Pharmakoepidemiologische Forschungsdatenbank

Das FDZ GePaRD basiert auf Daten gesetzlicher Krankenversicherungen in Deutschland seit 2004. Mit GePaRD können u. a. Forschungsfragen zum Gebrauch und zur Sicherheit von Arzneimitteln und Impfstoffen in der Routineversorgung untersucht werden, sofern die jeweilige Datennutzung gemäß § 75 SGB X genehmigt wird.

<https://www.bips-institut.de/forschung/forschungsinfrastrukturen/gepard.html>



#### FDZ GML



#### Forschungsdatenzentrum German Microdata Lab bei GESIS

Forschungsbasierter Service zu Mikrodaten der amtlichen Statistik in Europa und Deutschland: Tools für Datenmanagement und Datenanalyse. Metadaten (MISSY): Umfassende und detaillierte Datendokumentation amtlicher Mikrodaten. Wissenstransfer: Beratung, Workshops und Nutzerkonferenzen zur methodischen und inhaltlichen Forschungsfragen bei der Analyse amtlicher Mikrodaten. Gegründet 1987.

<https://www.gesis.org/gml/gml-home>



#### FDZ Internationale Umfrageprogramme



#### Forschungsdatenzentrum Internationale Umfrageprogramme bei GESIS

International vergleichende Umfragedaten aus über 70 Ländern der Welt zu fast allen Themen der Sozialwissenschaften: Comparative Study of Electoral Systems (CSES), European Values Study (EVS), Eurobarometer, Europa-Wahlen (EES), International Social Survey Programme (ISSP).

<https://www.gesis.org/angebot/daten-analysieren/internationale-umfragen>



#### FDZ IQB



#### Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen

Deutsche Datensätze aus den großen nationalen und internationalen Schulleistungsstudien sowie aus nationalen Studien mit Kompetenzmessungen im Bildungsbereich.

<https://www.iqb.hu-berlin.de/fdz>



#### FDZ-IWH

















#### Forschungsdatenzentrum des Leibniz-Instituts für Wirtschaftsforschung Halle

Unternehmensdaten zu Entwicklungstendenzen in der Industrie und im Baugewerbe Ostdeutschlands, zu Privatisierungsaktivitäten der Treuhandanstalt, zur Standortwahl multinationaler Unternehmen in den neuen Bundesländern und in Mittel- und Osteuropa sowie zu Produktivitäts- und Wettbewerbsfähigkeitsindikatoren europäischer Länder nach Firmengrößenklassen.

<https://www.iwh-halle.de/forschung/daten-und-analysen>



<b>FDZ IZA, IDSC</b> 	<b>Internationales Forschungsdatenzentrum des Forschungsinstituts zur Zukunft der Arbeit</b> <p>Nationale und internationale Arbeitsmarktdatensätze mit standardisierten Informationen (eddi-conferences.eu). Forschung mit sowie Methoden und Ressourcen für die Nutzung von Onlinedaten in Arbeitsmarktökonomie und Sozialwissenschaften. Entwicklung von Werkzeugen und Methoden für Remote Access (statsdirect.org) und Remote Processing (JoSuA).</p> <p><a href="https://www.iza.org/forschung/idsc">https://www.iza.org/forschung/idsc</a></p>	
<b>FDZ im KBA</b> 	<b>Forschungsdatenzentrum im Kraftfahrt-Bundesamt</b> <p>Das Forschungsdatenzentrum im Kraftfahrt-Bundesamt (FDZ im KBA) bietet für die wissenschaftliche Forschung anonymisierte Mikrodaten zum Zugang in das Fahreignungsregister (FAER) als Scientific Use Files an.</p> <p><a href="https://www.kba.de/DE/Statistik/Forschungsdatenzentrum/forschungsdatenzentrum_node.html">https://www.kba.de/DE/Statistik/Forschungsdatenzentrum/forschungsdatenzentrum_node.html</a></p>	
<b>FDZ-LifBi</b> 	<b>Forschungsdatenzentrum des Leibniz-Instituts für Bildungs- verläufe e.V. an der Otto-Friedrich-Universität Bamberg</b> <p>Längsschnittdaten des seit 2010 laufenden Nationalen Bildungspanels (NEPS - National Educational Panel Study) mit über 60.000 Teilnehmenden in sechs Startkohorten zur Erforschung von Kompetenzentwicklungen, Bildungsprozessen, Bildungsentscheidungen und Bildungsrenditen in formalen, nicht-formalen und informellen Kontexten.</p> <p><a href="https://www.lifbi.de/Institut/Organisation/Forschungsdatenzentrum-LifBi">https://www.lifbi.de/Institut/Organisation/Forschungsdatenzentrum-LifBi</a></p>	
<b>FDZ-pairfam</b> 	<b>Forschungsdatenzentrum des Beziehungs- und Familienpanels</b> <p>Datensätze aus der repräsentativen, interdisziplinären Längsschnittstudie zu partnerschaftlichen und familialen Lebensformen in Deutschland: „Panel Analysis of Intimate Relationships and Family Dynamics“ (pairfam).</p> <p><a href="https://www.pairfam.de">https://www.pairfam.de</a></p>	
<b>FDZ PIAAC</b> 	<b>Forschungsdatenzentrum PIAAC bei GESIS</b> <p>Nationale und internationale Datensätze zu PIAAC (Programme for the Assessment of Adult Competencies). Für Deutschland stehen des Weiteren Regionaldaten sowie Längsschnittdaten zur Verfügung.</p> <p><a href="https://www.gesis.org/piaac/fdz">https://www.gesis.org/piaac/fdz</a></p>	
<b>FDZ PsychData des ZPID</b> 	<b>Forschungsdatenzentrum PsychData des Leibniz-Zentrums für Psychologische Information und Dokumentation</b> <p>Bündelung quantitativer Datensätze sowohl aus der Grundlagenforschung als auch aus der angewandten Psychologie, Ausrichtung des Sammelauftrages insbesondere auf Längsschnittstudien, groß angelegte Umfragestudien und Testentwicklungen.</p> <p><a href="https://www.psychdata.de">https://www.psychdata.de</a></p>	
<b>FDZ Qualiservice</b> 	<b>Forschungsdatenzentrum Qualiservice</b> <p>Qualiservice bietet die Archivierung, Kuration und Bereitstellung qualitativer Forschungsdaten unterschiedlicher Disziplinen an. Die Services sind sicher, flexibel und forschungsorientiert und beinhalten umfassende Beratung und Aufbereitung von qualitativen Primärstudien für die Nachnutzung, Langzeitarchivierung, die Bereitstellung archivierter Forschungsdaten und relevanter Kontextinformationen.</p> <p><a href="https://www.qualiservice.org">https://www.qualiservice.org</a></p>	



Datenangebot: ■ Sozial ■ Wirtschaft ■ Bildung ■ Gesundheit ■ Psychologie ■ Qualitativ ■ Weitere

#### FDZ RKI



#### Forschungsdatenzentrum des Robert Koch-Instituts

ROBERT KOCH INSTITUT



Daten zum Gesundheitszustand und zum Gesundheitsverhalten der Bevölkerung, die auf der Basis von bundesweit repräsentativen Studien gewonnen wurden.

<https://www.rki.de/puf>

#### FDZ Ruhr am RWI



#### Forschungsdatenzentrum Ruhr am RWI – Leibniz-Institut für Wirtschaftsforschung



Spezialisierung auf regionale Daten: Angebot von Daten auf Ein-Quadratkilometer-Rastern. Neben der Geo-Referenzierung von Daten auf wissenschaftlicher Basis werden unterschiedliche Individual- und Unternehmensdaten aus Forschungsprojekten des RWI zur Verfügung gestellt.

<https://www.rwi-essen.de/forschung-und-beratung/fdz-ruhr>

#### FDZ-RV



#### Forschungsdatenzentrum der Rentenversicherung



Daten zu Versicherungskonten der in der gesetzlichen Rentenversicherung Versicherten, in denen der Versicherungsverlauf sowie Leistungen an die Versicherten aus Rente und Rehabilitation gespeichert sind.

<http://www.fdz-rv.de/FdzPortalWeb>

#### FDZ SFB 882 (2016 aufgelöst)

Qualitative und quantitative Datensätze aus dem Bereich der Ungleichheitsforschung.



(Das FDZ wurde 2016 aufgelöst, je nach Art und Grundlage wurden die Daten des FDZ SFB 882 an verschiedene Organisationen weitergegeben: IAB Daten gingen an das FDZ BA im IAB; qualitative Daten mit Organisationsbezug gingen an das FDZ BO; die restlichen Daten gingen an das fakultätsinterne Archiv SOBI an der Universität Bielefeld (im Aufbau befindlich). (Stand 14. September 2017)

<https://sfb882.uni-bielefeld.de/de/fdz-sfb882.html>

#### FDZ-SHARE



#### Forschungsdatenzentrum des Survey of Health, Ageing and Retirement in Europe



Daten der multidisziplinären Längsschnittstudie „SHARE“, die in bislang sieben Wellen Daten von ca. 140.000 Individuen zu Gesundheit, sozio-ökonomischem Status sowie sozialen und familiären Netzwerken von Personen ab 50 Jahren in über 20 europäischen Ländern und in Israel erhoben hat. Die 8. Welle von SHARE wurde 2020 erhoben.

<http://www.share-project.org/data-access.html>

#### FDZ SOEP



#### Forschungsdatenzentrum des Sozio-ökonomischen Panels am DIW Berlin



Daten aus repräsentativ erhobenen Wiederholungsbefragungen von Privathaushalten. Die SOEP-CORE Stichprobe umfasst Themen wie Einkommen, Erwerbstätigkeit, Bildung und Gesundheit. Zudem gibt es die längsschnittliche Innovationsstichprobe (SOEP-IS), die es externen Forschenden ermöglicht, eigene Forschungsvorhaben einzubringen.

<https://www.diw.de/SOEPfdz>

**FDZ Wahlen****Forschungsdatenzentrum Wahlen bei GESIS**

Zugang zu nationalen Wahlstudien (Bundes- und Landtagswahlen), Politbarometer, Forsa-Bus, ARD Deutschlandtrend, Umfragen der Bundesregierung. Das derzeit größte Projekt im FDZ Wahlen ist die German Longitudinal Election Study (GLES).

<https://www.gesis.org/wahlen>

**FDZ  
Wissenschafts-  
statistik****Forschungsdatenzentrum Wissenschaftsstatistik des  
Stifterverbandes**

Daten über die Forschungs- und Entwicklungsaktivitäten der deutschen Wirtschaft, zum finanziellen Volumen, der Struktur und der regionalen Verteilung von Forschung und Entwicklung (FuE) sowie der in diesem Bereich tätigen Personen.

<https://www.fdz-wissenschaftsstatistik.de>

**IÖR-Monitor**

(vorläufig  
akkreditiert)

**Monitor der Siedlungs- und Freiraumentwicklung**

Der IÖR-Monitor ist eine Forschungsdateninfrastruktur des Leibniz-Instituts für ökologische Raumentwicklung (IÖR). Er stellt Daten und Informationen in hoher räumlicher Auflösung zur Flächennutzungsstruktur und deren Entwicklung sowie zur Landschaftsqualität für die Bundesrepublik Deutschland bereit.

<https://www.ioer-monitor.de>

**SAFE FDZ**

(vorläufig  
akkreditiert)

**Forschungsdatenzentrum des Leibniz-Institut für Finanz-  
marktforschung SAFE**

Der Mangel an pan-europäischen Finanzdaten führt dazu, dass Forscher auf US Daten ausweichen und Forschungsergebnisse nur eingeschränkt auf den europäischen Raum übertragen können. Das SAFE Forschungsdatenzentrum bündelt nicht nur vorhandene Daten, sondern erhebt und erstellt auch neue deutsche und europäische Datensätze, um die europäische Perspektive der empirischen Forschung zu stärken.

<https://safe-frankfurt.de/de/data-center.html>

**ZEW-FDZ****Forschungsdatenzentrum des Zentrums für Europäische  
Wirtschaftsforschung**

Das ZEW-FDZ stellt die Individualdaten von ZEW-Firmenbefragungen zu Innovationsaktivitäten, Entwicklung junger Unternehmen, Nutzung von Informations- und Kommunikationstechnologien und anderen Themen zur Verfügung. Ferner können die Daten einer ZEW-Expertenbefragung, der „ZEW-Finanzmarkttest“, genutzt werden.

<https://www.zew.de/zew-fdz>

**FDZ WSI**

(Gaststatus im  
FDI Ausschuss)

**Forschungsdatenzentrum der Hans-Böckler Stiftung am WSI**

Das FDZ WSI bietet Zugang zu den WSI-Betriebsrätebefragungen bis 2011.

<https://www.wsi-datenzentrum.de>



## Anhang C

### Die Monitoringkommission

Für die Qualitätssicherung hat der RatSWD die Etablierung einer Monitoringkommission ab Juli 2016 beschlossen. Ihre Hauptaufgabe ist das Sammeln und Bewerten der regelmäßigen Berichte der FDZ. Ferner überprüft die Kommission die Erfüllung der Auflagen im Fall von vorläufigen Akkreditierungen. Die Kommission wird für je drei Jahre zeitlich parallel zur RatSWD-Legislaturperiode vom FDI Ausschuss aus den eigenen Reihen gewählt und genießt damit besonderes Vertrauen und Legitimität. Sie besteht aus vier Mitgliedern des FDI Ausschusses und zwei Stellvertretenden (für den Fall der Verhinderung eines gewählten Mitglieds) sowie den Vorsitzenden des RatSWD als Gästen.

#### Mitglieder der Monitoringkommission

**Maurice Brandt** (Vorsitz Mai 2019 – 31. Dezember 2019)

Forschungsdatenzentrum des Statistischen Bundesamtes (FDZ-Bund)

**Dr. Lea Eilers** (Vorsitz Juli 2018 – September 2018)

Forschungsdatenzentrum Ruhr am RWI – Leibniz-Institut für Wirtschaftsforschung (FDZ Ruhr am RWI)

**Dr. Cornelia Lang** (Vorsitz seit 01.01.2020)

Forschungsdatenzentrum des Leibniz-Instituts für Wirtschaftsforschung Halle (FDZ-IWH)

**Holger Quellenberg** (Vorsitz Oktober 2018 – April 2019)

Forschungsdatenzentrum des Deutschen Jugendinstituts (FDZ-DJI)

**Dr. Sandra Gottschalk** (Vertreterin)

ZEW-Forschungsdatenzentrums (ZEW-FDZ)

**Tatjana Mika** (Vertreterin)

Forschungsdatenzentrum der Rentenversicherung (FDZ-RV)

#### Ständige Gäste der Monitoringkommission

**Prof. Stefan Bender**

Stellvertretender Vorsitzender des RatSWD (von Juli 2014 bis November 2019)

**Prof. Regina T. Riphahn, Ph.D.**

Vorsitzende des RatSWD (von Juli 2014 bis Juli 2020)

**Prof. Dr. Jürgen Schupp**

Stellvertretender Vorsitzender des RatSWD (von November 2019 bis Juli 2020)

**Prof. Dr. Monika Jungbauer-Gans**

Vorsitzende des RatSWD (seit Juli 2020)

**Prof. Dr. Kerstin Schneider**

Stellvertretende Vorsitzende des RatSWD (seit Juli 2020)

## Anhang D

### Mitwirkende bei der Erstellung des Tätigkeitsberichts 2019

**Maurice Brandt** (Mitglied der Monitoringkommission)

Forschungsdatenzentrum des Statistischen Bundesamtes (FDZ-Bund)

**Dr. Mathias Bug** (Geschäftsstelle RatSWD)

**Dr. Lea Eilers** (Mitglied der Monitoringkommission)

Forschungsdatenzentrum Ruhr am RWI – Leibniz-Institut für Wirtschaftsforschung (FDZ Ruhr am RWI)

**Dr. Anna Fräßdorf** (Geschäftsstelle RatSWD)

**Dr. Sandra Gottschalk** (Stellvertretendes Mitglied der Monitoringkommission)

ZEW-Forschungsdatenzentrums (ZEW-FDZ)

**Dr. Cornelia Lang** (Vorsitzende der Monitoringkommission)

Forschungsdatenzentrum des Leibniz-Instituts für Wirtschaftsforschung Halle (FDZ-IWH)

**Tatjana Mika** (Stellvertretendes Mitglied der Monitoringkommission)

Forschungsdatenzentrum der Rentenversicherung (FDZ-RV)

**Holger Quellenberg** (Mitglied der Monitoringkommission)

Forschungsdatenzentrum des Deutschen Jugendinstituts (FDZ-DJI)

**Dr. Katrin Schaar** (Geschäftsstelle RatSWD)







## Impressum

**Herausgeber:**

Rat für Sozial- und Wirtschaftsdaten (RatSWD)

Rungestr. 9

10179 Berlin

office@ratswd.de

<https://www.ratswd.de>

**Redaktion:**

Dr. Katrin Schaar, Dr. Mathias Bug

**Gestaltung/Satz:**

Claudia Kreutz

**Icons:**

made by Freepik from <https://www.flaticon.com>

Berlin, Oktober 2020

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) unter dem Förderkennzeichen 01UW1802 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt, sofern nicht anders ausgewiesen, beim RatSWD.

doi: [10.17620/02671.56](https://doi.org/10.17620/02671.56)

**Zitationsvorschlag:**

RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2020): Tätigkeitsbericht 2019 der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). <https://doi.org/10.17620/02671.56>



■ **Der Rat für Sozial- und Wirtschaftsdaten (RatSWD)** berät seit 2004 die Bundesregierung und die Regierungen der Länder in Fragen der Forschungsdateninfrastruktur für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften. Im RatSWD arbeiten zehn durch Wahl legitimierte Vertreterinnen und Vertreter der sozial-, verhaltens- und wirtschaftswissenschaftlichen Fachdisziplinen mit zehn Vertreterinnen und Vertretern der wichtigsten Datenproduzenten zusammen.

Er versteht sich als institutionalisiertes Forum des Dialoges zwischen Wissenschaft und Datenproduzenten und erarbeitet Empfehlungen und Stellungnahmen. Der RatSWD engagiert sich für eine Infrastruktur, die der Wissenschaft einen breiten, flexiblen und sicheren Datenzugang ermöglicht. Solche Daten werden von staatlichen, wissenschaftsgetragenen und privatwirtschaftlichen Akteuren bereitgestellt. Der RatSWD hat 38 Forschungsdatenzentren (Stand: Mai 22020) akkreditiert, deren Kooperationen er fördert.



[www.ratswd.de](http://www.ratswd.de)